The Journal of
APPLIED PSYCHOLOGY

EDITED BY

JAMES P. PORTER
Clark University

WILLIAM F. BOOK
Indiana University

And a Board of Co-operating Editors

VOLUME IV

1920

WORCESTER, MASS.
FLORENCE CHANDLER, Publisher
Reprinted with the permission of the original publishers.
JOHNSON REPRINT CORPORATION
KRAUS REPRINT CORPORATION

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. V. Bingham, Carnegie Institute of Technology; June E. Downey, *University of Wyoming*; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; Joseph Jastrow, University of Wisconsin; Herbert S. Langfeld, Harvard University; James H. Leuba, Bryn Mawr College; Henry C. McComas, Princeton University; Walter D. *Scott, Northwestern* University; C. E. Seashore, University of Iowa; Daniel Starch, University of Wisconsin; Edward K. Strong, Jr., Carnegie Institute of Technology, Pittsburgh; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; J. E. Wallace Wallin, Psycho-Educational Clinic, St. Louis; G. M. Whipple, University of Michigan; Helen Thompson Woolley, Vocation Bureau, Cincinnati.

The scope of the Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and *problems of business appeal. (b) Studies of individual mentalities, such* as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis (c) The influence of general environmental conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, *etc. (d) The psychology of everyday activities, such as reading, writing,* speaking, singing, playing games or musical instruments, sports, etc.

Copyright, 1919, by G. Stanley Hall.

TABLE OF CONTENTS

HENRY FOSTER ADAMS
- The effect of climax and anticlimax order of presentation on memory 330-338

JOHN E. ANDERSON
- A comparison of two methods of giving the number series completion test 346-347

G. F. ARPS and H. E. BURTT
- Correlation of army alpha intelligence test with academic grades in high schools and military academies . . 289-293

H. L. BAKER and EDNA GORDON
- Intelligence tests and academic standing 361-363

MILDRED FRANCES BAXTER
- Opportunities for college graduates in psychological examining in social service work and education . . 207-218

M. F. BEESON
- Intelligence at senescence 219-234

JAMES W. BRIDGES and ESTHER REEDY
- A short point scale for mental measurement . . . 258-262

HAROLD E. BURTT
- Employment psychology in the rubber industry . . 1-17

H. E. BURTT and G. F. ARPS
- Correlation of army alpha intelligence test with academic grades in high schools and military academies . 289-293

HIRAM BYRD
- A case of phenomenal memorizing by a feeble-minded negro 202-206

O. R. CHAMBERS and S. L. PRESSEY
- First revision of a group scale designed for investigating the emotions, with tentative norms 97-104

S. M. DERRICK
- A comparative study of the intelligence of seventy-five white and fifty-five colored college students by the Stanford revision of the Binet-Simon scale . . 316-329

EDGAR A. DOLL
- The degree of Ph.D. and clinical psychology. . . 88-90

JOSEPHINE CURTIS FOSTER
- Significant responses in certain memory tests . . 142-154

JOSEPHINE CURTIS FOSTER and GRACE A. TAYLOR
- The applicability of mental tests to persons over fifty years of age 39-58

FRANK N. FREEMAN
- Clinical study as a method in experimental education . 126-141

ELLIOT FROST
- What industry wants and does not want from the psychologist 18-24
- Should psychology bake bread 294-305

THOMAS R. GARTH
- Racial differences in mental fatigue 235-244

ROBERT H. GAULT
- Picture completion 310-315

SYBIL MARSH GIBSON
- A decision study of 150 young men and women . . . 364-374

EDNA GORDON and H. J. BAKER
- Intelligence tests and academic standing . . . 361-363

CONTENTS

MARIE GROETZINGER and A. A. ROBACK
The applied psychology of names 348-360
K. W. IRLE and S. C. KOHS
Prophesying army promotion 73-87
S. C. KOHS and K. W. IRLE
Prophesying army promotion 73-87
HENRY C. LINK
A new application of psychology to industry . . 245-249
A further development of employment psychology . 306-309
LAWRENCE MARCUS
Vocational selection for specialized tasks 186-201
BRUCE V. MOORE
General intelligence determined by its weakest essential element 155-161
E. LEIGH MUDGE
The common synaesthesia of music 342-345
ELSIE MURRAY
Psychological tests as diagnostic of vocational aptitudes in college women 30-38
ARTHUR S. OTIS
The selection of mill workers by mental tests . . 339-341
JOSEPH PETERSON
Tentative norms in the rational learning test . . 250-257
RUDOLPH PINTNER and SAMUEL RENSHAW
A standardization and weighting of two hundred analogies 263-273
LUELLA WINIFRED PRESSEY
The influence of (a) inadequate schooling and (b) poor environment upon results with tests of intelligence . 91-96
S. L. PRESSEY and O. R. CHAMBERS
First revision of a group scale designed for investigating the emotions, with tentative norms 97-104
ESTHER REEDY and JAMES W. BRIDGES
A short point scale for mental measurement . . 258-262
SAMUEL RENSHAW and RUDOLPH PINTNER
A standardization and weighting of two hundred analogies 263-273
A. A. ROBACK and MARIE GROETZINGER
The applied psychology of names 348-360
GRACE A. TAYLOR and JOSEPHINE CURTIS FOSTER
The applicability of mental tests to persons over fifty years of age 39-58
EDWARD L. THORNDIKE
A constant error in psychological ratings . . . 25-29
Equality in difficulty of alternative intelligence examinations 283-288
M. J. VAN WAGENEN
Some results and inferences derived from the use of the army tests at the University of Minnesota . . . 59-72
CHARLES H. WOOLBERT
Effects of various modes of public reading 162-185
Tables to facilitate the computation of coefficients of correlation by the rank difference method 115-125
NOTES 112-114
BOOK REVIEWS 105-111, 274-282, 375-376

THE
JOURNAL OF APPLIED PSYCHOLOGY

VOL. IV MARCH, 1920 No. 1

EMPLOYMENT PSYCHOLOGY IN THE RUBBER INDUSTRY

By HAROLD E. BURTT, Ohio State University

1. *Introduction*

The writer spent several months in the spring of 1919 as Consulting Psychologist of a large Canadian rubber company. Most of the time was devoted to the development of tests for the use of the employment office of a rubber tire factory. It was not originally intended to publish an account of this work, because it involved no radically new kinds of tests, because there were no inter-correlations of theoretical interest and because if detailed test procedure was published it might be used by the incompetent or might even reach the hands of prospective employees. However, after reading Link's recent book on *Employment Psychology,* it seemed worth while to give a general account of the methods used at the tire factory, as they illustrate the solution of problems similar to those recounted by Link, but with a quite different statistical approach. In both instances the general principle was to correlate test scores with known vocational ability, but in one case a wide range of factory and clerical operations was covered with rather simple statistical treatment and in the present instance fewer operations were studied, but with the use of partial correlations. Furthermore the personal experiences of the writer in methods of organizing the work, enlisting the co-operation of executives, and arranging for the continuance of the work might be of interest. In the near future when this field is quite apt to be exploited by those with too little ability and too much enthusiasm, there will be a need for a restraining influence in the form of conservative accounts of personnel work done with strictly scientific procedure. It is hoped that the following will be a slight contribution of this sort.

2. *Preliminary Work*

The preliminary work at the factory fell into two parts,—getting personally oriented and getting in the proper rapport with those in authority. It was necessary at the outset to become familiar with the different types of work carried on at the factory and to learn the technical terminology. This knowledge was necessary for later analysis of the mental factors involved in various factory operations. One of the assistants to the Factory Superintendent was detailed as aide to the writer,—conducting him about the factory, explaining the work and introducing him to various foremen so that it was possible to get acquainted and talk over the mental and physical qualifications for various kinds of work. A further study was made of the methods of accounting, time study and evaluating piece-work. These facts were of value in getting production ratings of men who were tested, and following up new workers. Data were secured on labor turnover to see what departments apparently needed the most attention. It was necessary, also, to become familiar with the executive organization in order to know to whom to go for information or to get various things done.

The second preliminary step was to get in the proper rapport with the personnel whose co-operation would be needed in the work. The factory at the time was managed by an operating council of five men, rather than a single manager. An early meeting with this council was arranged in order to discuss plans. These men readily saw the purpose and the general methods of the work. Shortly after this a meeting of all the foremen was called. They had an organization of their own and this meeting was scheduled as one of their regular meetings with the writer as an outside speaker. A talk was given about personnel work in the army leading up to the question of round pegs in square holes in industry. The matter was finally brought home to them and the proposed work explained in detail and their co-operation enlisted. Practically all of them were much interested in the matter and this preliminary discussion was all that was ever necessary in obtaining their co-operation in the work. This rapport was also furthered by the writer's frequent presence in the factory. He visited different departments very often, talked with the foremen, watched the workmen and made himself felt generally as a part of the organization. Shortly after the foremen's meeting a preliminary series of tests covering a rather wide range of mental factors was given to the executives, to most of the foremen, and to a sampling of the factory workers. This served the purpose of familiarizing them with the nature of mental tests. This was especially desirable because from information that reached the writer indirectly it was found that the conception of mental tests held by the majority of the factory personnel varied from phrenology to microscopic x-ray examination of the brain. The results of this preliminary series were of some statistical import, but their greatest function lay in familiarizing those whose co-operation was especially needed with the methods.

3. *Obtaining the Criterion*

As early as possible in the research (as soon as the foremen had been tested) vocational ratings of a large number of workers in the factory were obtained. The writer learned a bitter lesson in the Air Service in testing large numbers of men and being unable to get subsequent vocational ratings. It was highly advantageous to have the criterion before any tests were given at all. This made it possible to send to the factory the names of men whom it was desired to test. Thus when only a sampling of those engaged in a given type of work was to be tested, it was feasible to get a sampling corresponding roughly to a normal distribution curve for ability in that work. Otherwise the sampling might comprise those at one extreme of ability or only those in the middle range with resulting lower correlations due to the homogeneity of the group. Furthermore if some of the men desired were unavailable (e. g., were on the night shift) it was possible to substitute others of similar vocational ability.

Two measures of vocational ability were desired in order to correlate with two scores in each test, and correct the coefficient for attenuation. In several operations three ratings besides the production ratings (piece-work) were available; the inspector who had perhaps a dozen men under him furnished one rating, the foreman who had charge of finishing or building or cord building, etc., furnished a second, and the head foreman of the whole floor a third. One of these could then be combined with the production rating and the other two averaged.

The rating sheet used was similar to that described by Miner.[1] Five adjacent columns each ten millimeters in width were headed "high test 5th," "middle 5th," etc. The names of the workmen were typed in a column at the left and the foremen requested to place after each man's name a cross in the appropriate column. The foremen were further encouraged to grade a man's ability as finely as possible by placing the cross to the right or to the left of the column. These ratings were measured in millimeters from the edge of the chart and furnished a score between 0 and 100. Inasmuch as the ratings by different foremen and inspectors varied considerably,—some rating all their men higher than did others,—each foreman's ratings were averaged, the standard deviation computed and the original measures converted into ratios of individual deviation to standard deviation for that foreman. The production ratings were treated in the same manner. It was then valid to average one foreman's rating with the piecework score and to average together the other two foremen's ratings. These two figures could then be correlated with the two test measures (infra). This method was used in all correlations where the coefficient was corrected for attenuation. In some factory operations which were not to any considerable extent specialized or in which only a small number of workers were engaged, the foremen were asked merely to select a few of the best and a few of the worst workmen. The desirability of correcting the coefficient of correlation whenever possible was shown by the low correlations sometimes found between the ratings of the same workmen by different foremen. These coefficients were sometimes as low as .60 whereas correlations between two parts of the same test were generally in the vicinity of .90.

4. *Selection of Tests*

There are two current tendencies in developing methods of vocational selection. The first strives to reproduce in miniature the work or the mental situation involved. This method might be illustrated by Dodge's tests for Gun-pointers in the Navy. The second attempts to analyze the mental abilities involved in the work, and test these separately. This method might be typified by the Air Service tests of aptitude for flying. In both cases the procedure is, of course, empirical,—evaluating the methods upon persons of known ability before applying them to unknown applicants.

---

[1] J. B. Miner. "Evaluation of a Method for Finely Graded Estimates of Abilities." *Journal of Applied Psychology*, 1917, 1, 123-133.

The latter of these two methods was used in the present instance. The specialized forms of work in a tire factory involved so many minor operations as to necessitate very complex apparatus. Several involved the handling of new rubber stock, which is sticky,—a condition difficult to reproduce in standardized form in the laboratory. Consequently various forms of work were analyzed into their mental and motor components, and a limited number of tests measuring these components were applied to workers of a given sort.

The majority of the tests were given by means of mimeographed blanks to a group of people simultaneously. Others involved individual examination and a certain amount of apparatus. Thirty-two tests were used in all covering a rather wide range. It is difficult and unnecessary to say what mental factor a test specifically measures, but there were in the list of tests selected some in which the emphasis lay upon motor processes, reaction time, attention, observation, learning, association, memory, reasoning, space and time perception, social insight, ingenuity and ability to follow directions.

The tests given on blanks were adaptations of recently devised methods drawn from various sources, such as the material left by the late Professor Münsterberg, the work of the Carnegie Institute of Technology, various test blanks used in the Air Service and the Sanitary Corps in the army, and scattered periodical accounts. In most cases the material was worked over considerably. Anything distinctly academic in categories, terminology or range of information was simplified to come within the range of the average workman. Geographical factors were likewise adapted to local conditions and all extraneous matters eliminated as far as possible. Each test was given on a separate sheet with a brief example at the top. Standardized instructions involving the explanation of the example were used. Tests in this form were suitable for group examination. They were all performed with a time limit. They were tried first of all upon a few people in order to determine a proper time interval which would not be unnecessarily long but would still give considerable scatter in amount done by different persons and in order to include enough material so that no one would finish the test within the allotted time. The usual precautions were taken as to seating, lighting, sharpening of pencils and starting and stopping at signal.

The individual tests were mostly adaptations of the conventional motor tests, described in the manuals such as Whipple's. Various types of reaction time were also measured with a chronoscope. A few tests devised by the psychology sub-committee on Aviation of the National Research Council were included. A few people were given these tests as a preliminary step to determine the number of trials or the rate of performance (if done at a controlled speed). The amount of practice and the instruction for each performance were standardized.

All the tests except those in the preliminary series described below, in which no careful correlations were to be obtained, were given in two installments. Ideally these should occur on separate days, but it was fairly satisfactory to go through the first installments of each test given to a particular person or group of persons in turn and then go through the second installments in the same order. Each installment was divided in two equal parts on the basis of time. This was facilitated in group examination by having the persons make a check mark at the point where they were working at the given signal. With these four test measures the first one of the first installment was averaged with the last one of the last installment and the other two likewise averaged together. This gave two test measures which could be correlated with the two vocational measures described above to correct the coefficient of correlation for attenuation.

5. *The Laboratory*

All the testing in the research stages of the work was done by the writer. A large well lighted room in the office building was equipped for a laboratory, containing the apparatus for the individual tests and table facilities for testing several people simultaneously on group tests. For larger groups (ten or twelve) a large private office with extra tables was used.

It is sometimes suggested that a laboratory of this sort should be in the factory itself, in order to obtain the workmen at the right time, to afford no sharp contrast with shop conditions and consequent nervousness, and to keep the experimenter in touch with the shop work. In the present instance the use of the laboratory was perfectly satisfactory. A list was sent the day before to the man in charge of time cards at the factory of the men whom it was desired to test on the morrow, together with the hour at which they were wanted and the approximate length of time they would be needed. There appeared to be no great shock on the part of the workmen in coming from the factory to the more quiet laboratory, for the news of the work and the nature of the tests spread rapidly through the personnel. Furthermore the writer found ample contact with factory conditions in interviewing foremen and watching various men at work.

6. *Preliminary Series*

The first actual testing done at the factory consisted of a set of twenty group tests covering a fairly wide range of mental factors and occupying about two hours. The purposes of this series were three: first,—as indicated above, to familiarize those in authority, especially the operating council and foremen with the methods so that they would be in a better position to co-operate in getting workmen to take the tests and in providing ratings; secondly,—to determine which types of tests were most promising for further study in connection with special factory operations; thirdly,—to obtain a rough notion of the distribution of intelligence throughout the factory.

The tests were given to the operating council, other executives in the factory, some of the clerical force and laboratory personnel and a sampling of workers in a considerable number of factory operations, including equal numbers of good and poor workers of each sort. After the foremen had been tested there was a pause while getting ratings of the workmen in order to select the aforementioned samplings.

Each test was averaged for each occupational group and also for the entire number of people tested. The group averages, and in some cases individual scores, were all tabulated as a ratio of the score in a test to the grand average for that test. It was found that occupational group averages in some tests showed a much greater difference between extremes than in other tests. For example, the extremes of one test ranged from 81 to 112 per cent, whereas in another test they ranged from 5 to 184 per cent. The first of these would obviously be of little value in differentiating groups of workers of the sort studied. Furthermore certain groups were markedly superior or inferior in certain tests. These facts together with a detailed individual study of good versus poor workers in various groups, gave the starting point for subsequent work, indicating, for example, whether motor, attentive or memory processes were most involved in a given factory operation.

All twenty tests were averaged together for each occupational group. This average of the twenty tests, inasmuch as they covered a fairly wide range, gave a rough index of general intelligence. Presumably any combined score of a considerable range of tests gives such an index and it is merely then a question of determining norms. This is what is done, of course, in various point-scales for intelligence. Comparing the different groups of workers in intelligence there was found an occupational hierarchy. The operating council, laboratory force and clerical workers stood rather high in intelligence. The shipping clerks and general factory executives were appreciably lower. Still lower came the foremen and fairly skilled workers, such as finishers and builders. Below these were the less specialized forms of factory workers and at the bottom of the scale the unskilled laborers. This occupational hierarchy is interesting in its agreement with the similar one found in the army on the basis of the Alpha examination.[2] It would seem that there are different kinds of industrial performance, each with its minimum intelligence requirements and that a person tends finally to reach the highest level for which his intelligence qualifies him.

7 *Statistical Methods*

In subsequent series test scores were correlated with known vocational ability, in order to select the best tests and weight them properly to obtain the best possible vocational prediction.

As soon as the ratings and the test scores were in final form, the correlation formula involving rank differences squared was used for a preliminary selection of the tests. The best ones were then correlated by the "products-moments" formula. There were two scores in each test and two measures of vocational ability as above described. The necessary correlations of these were computed to be used in the following formula:

$$\frac{\sqrt[4]{V_1T_1 . V_1T_2 . V_2T_1 . V_2T_2}}{\sqrt{V_1V_2 . T_1T_2}}$$

where V_1 and V_2 are the two vocational ratings (cf. section 3 supra) and T_1 and T_2 are the two test measures, (section 4 supra) and $V_1 T_1$ indicates the correlation between V_1 and T_1, etc. To correct the coefficients for the inter-test correlations, the simpler formula was used:

$$\frac{\sqrt{A_1B_2 . A_2B_1}}{\sqrt{A_1A_2 . B_1B_2}}$$

in which A_1 is the first measure in the test A, A_2 the second measure in test A and similarly for B_1 and B_2 and A_1B_2 indicates the correlation between these two measures. This latter formula seemed justifiable for inter-test correlations because the two parts of each test correlate highly, whereas foremen's ratings do not inter-correlate as highly, and consequently more careful correction is necessary.

---

[2] The only published account of this study is an abstract of a paper read at the 1918 meeting of the American Psychological Association by J. W. Bridges, *Psychological Bulletin*, 16, 1919, p. 42.

The coefficients of partial correlation were then obtained for the given factory operation and the regression equation computed. The writer usually follows the method outlined by Rosenow.[3] The practical problem is to select as few tests as possible while still giving a high prediction. Five tests may give, for example, practically as good a prediction as seven tests with a consequent saving of time. Rosenow's method was followed of computing from a limited number of partial coefficients the final R of multiple correlation, i. e., the correlation of combined weighted test scores with the criterion. It is convenient in numbering the variables to take as the end ones, those most apt to be dropped,—for example with a criterion X_1 and 6 tests, X_2 to X_7, it is wise to select as X_2 and X_7 the tests which have lowest correlation with the criterion. Thus when the R of multiple correlation for all six tests has been found, without additional labor one can determine the R that will be obtained by dropping test two or test seven.

The regression equation was then worked out for the variables which were retained. A table for weighting the test scores was made, the original test measures weighted accordingly and the combined score correlated with the criterion as a check on the whole procedure. The individual combined weighted scores were now averaged and their standard deviation computed in order to make from the probability integral a decile distribution. A ten by ten fold table was then constructed from the probability integral for the given R of multiple correlation showing what per cent of those in the highest tenth in one variable were in the highest, second, third, etc., tenth in the other variable. Substituting for the deciles in one variable the deciles of combined weighted test scores and letting the other variable represent vocational ability, it was possible from the table to predict the probability of a person attaining a given score in the test of falling within any decile of vocational ability.

8. *Tests for Special Factory Operations*

It was necessary, of course, to limit the number of tests given to special operatives to those forms which had greatest promise. This limiting was done in two ways. In the first place the operations were analyzed as far as possible by observing men at work and talking with foremen as to the mental qualifications needed for the given operation. Secondly the preliminary series (supra) in which a few good and poor workers of various sorts took a wide range of tests, made it possible to limit the number of tests to the general categories which appeared of significance in this preliminary study.

---

[3] Rosenow, C. The Analysis of Mental Functions, *Psychological Review Monographs*, No. 106, 1917 (appendix).

The first operation studied intensively was that of *finishing tires*. There was a large labor turnover in this department and it apparently involved rather specialized ability. The worker received the tire partially built on the iron core, put stock of lighter weight on the outside and rolled it down, spinning the tire by hand. He was furnished with strips of stock of varying widths which had to be applied to certain parts of the tire, in some cases following a line rather closely. This operation was grouped with that of *treading tires* because the two were quite similar, except that in treading, no stock was applied near the rim. Thirteen tests were given to finishers of known ability,—tests involving motor ability, learning, attention, reaction time and judgment of distance and velocity. One surprising result was the low correlation of vocational ability with most of the tests involving motor coordination such as the conventional forms of tapping, aiming and steadiness tests. Six of the tests correlated rather highly with ability at the work, but it was found that three tests properly weighted gave practically as good a prediction. These tests involved underlining adjacent pairs of numbers whose sum was ten in a large group of numbers, finding consecutive numbers scattered at random over a page, and simple visual reaction. These tests properly weighted gave a correlation with finishing ability of .61.

The second group of operatives studied were those who *handed out stock*. The operation was fairly simple involving looking at the numbers on tags at the finishing benches, remembering the specifications long enough to go to the place where the stock was "booked" (placed between sheets of cloth to avoid sticking), and supplying the stock to the finisher. This was not a difficult operation, but had a considerable turnover. Ten tests were given involving memory, reasoning, directions, association, learning and attention. The regression equation for eight variables (that is seven tests) gave a correlation with ability at this work of .71. Using only three of these tests, a correlation of .67 could be obtained. These three tests involved underlining adjacent numbers whose sum was ten, underlining adjacent letters which formed words in an unseparated mixture of letters and memory tested by the method of word pairs.

The operation of *tire building* was somewhat like finishing. The iron core was placed on the machine, the stock from the rollers above pulled down and placed on the core, the clutch thrown with the foot to spin the tire and then the rollers applied by turning a small lever. Several plys were placed on each tire as well as a bead near the rim. Thirteen tests similar to those used with finishers were given to builders of known ability. The results were negative,—no correlation being greater than .16. There were three possible explanations: (1) None of the mental factors tested were involved in the building operation; (2) the vocational ratings were unreliable; (3) the men did not work with maximum effort at the tests. As to the first possibility it is doubtful, in view of the range of tests tried, that none of the factors involved in building were represented. As to the second point the correlation coefficients were corrected for attenuation, which should theoretically compensate, to a considerable extent, chance errors in the vocational estimates. The third explanation seemed the most probable. The average worker did not feel that his present job was permanent, but was considering possibilities of advancement or discharge. The average or good builder, however, received good pay and felt that his position was permanent and was consequently less interested in the tests (witness the fact that a number of builders refused to take the tests, whereas this seldom occurred with other operatives). The poor builders, however, might realize their ability and try hard in the tests. There were obvious cases where a builder appeared rather indifferent to the work. It was planned to check this factor later by testing new help who were hired for building. In the employment office they would doubtless work with maximum incentive.

The *clerical workers* were made the basis of a brief study. No effort was made to differentiate various kinds of office work, but the employees were rated merely for general clerical ability aside from stenography. Ratings were made by various managers and heads of departments. Ten tests involving attention, memory, association, and so-called alertness were used. Four of these tests were retained in the regression equation,—underlining adjacent letters which formed a word, reading a text in which alternate letters were to be omitted in order to make sense, an arithmetical test, and an analogies test in the form given in the army Alpha. This combined test score correlated with clerical ability to the extent of .56.

It was necessary to group together all the remaining *general factory operations*. In all the other forms of work there were either too few people engaged in the same sort of work to make a valid basis for statistics, or the work itself was quite simple and obviously did not need a specialized mental make-up. The problem then was to devise a simple group of tests to differentiate good workers in moderately skilled operations in general from poor workers, and to further differentiate this class from unskilled laborers. A sampling of good and poor workers was taken from various parts of the factory including final inspection, bicycle tires, bead wrapping, wire department, sundries, stock rolling and booking, solid tire, etc. A few of the best and worst workers of each sort were selected by the foremen. They were given nine different tests. With the ratings merely in two groups, detailed statistical treatment was impossible so the method of unlike signed pairs was used. Five tests were retained, inter-correlated and weighted by inspection. These tests were mostly of the sort generally characterized as intelligence tests such as memory for word pairs, association of opposites, absurdities, substitution and the like. There was a fairly close correspondence between combined test score and general ability in the factory. As nearly as could be estimated this correlation was about .50.

9. *Testing New Employees*

Most of the men passing through the employment office during about a month, with the exception of those who were obviously fitted for only unskilled labor, were tested. The ten by ten fold table described above was too complicated for use in the employment office, so a simplified form like the following was devised:

"Work; Tire Finishing

By good workers is meant the best three tenths of a large number of workers taken at random.

By average workers is meant the next three tenths, i. e., the 4th, 5th and 6th.

By poor workers is meant the next three tenths, i. e., the 7th, 8th and 9th.

By very poor workers is meant the lowest tenth.

Of a hundred men scoring in tests:

128 or better: 73 will be good, 21 average, 6 poor, 0 very poor.

103 to 127; 56 will be good, 31 average, 12 poor, 1 very poor, etc.

These figures instead of showing how many out of a hundred people obtaining a given score will be good, average, etc., may be used equally well to predict a probability of a single applicant; for example, if a man scores 128 or better, the chances are 73 out of 100 that he will be a good worker, 21 out of a 100 that he will be an average worker, etc."

It was desirable both in this blank and in personal conversation to emphasize the probability aspect of the prediction in order to minimize the effect of the occasional cases apparently violating the general rule. Business people are very apt to attach great weight to a single case in which a person who does well in the tests does poorly in the factory, neglecting the fact that there is always a certain probability of such a thing happening.

These tables for the various correlation coefficients or for the various factory operations were placed in the hands of the employment manager. Most of the recommendations at this time were made in person by the writer with these blanks before the manager. In a larger establishment where there were various parties concerned with hiring employees and where more occupations had been studied, it would doubtless be advisable to have a blank form printed for each factory operation, enter the individual's name on this blank, and place a check mark in the class into which he fell in his test scores, thus indicating at a glance his probable success in that work.

In the present instance effort was made to hire men who fell only in the best five or six tenths in test scores. The per cent of rejections would depend, of course, upon the labor market. During the war, for instance, industries were compelled to hire practically anyone who was available regardless of ability or probable success. Supposing, for example, the lowest ten per cent of workers of a given sort constituted the labor turnover, if men were hired at random for this work the turnover would be ten per cent. If test score and vocational ability correlated to the extent of .70 and enough applicants were available so that only the best half could be selected, the decile tables above described show that this labor turnover could be reduced to one per cent. With a coefficient of .60, the ten per cent turnover could be reduced to about two per cent. If the two lowest deciles represented labor turnover, with a coefficient of .70, this twenty per cent turnover could be reduced to about four per cent. Any other expectation on the basis of a given coefficient for a given percentage which constituted a turnover could be worked out from the tables.

The test program in the case of given workers depended largely upon the positions vacant in the factory. At the time most of the new help were tested by the writer, there was a considerable need for tire finishers. Consequently all applicants who appeared at all promising were given the tests described above for finishers. If a number presented themselves at once they were given together two of the tests and then each individual was taken separately through a measurement of visual reaction time. A stenographer provided with stencils for correcting the tests and tables for weighting them was able generally to keep up with this procedure, so that a few minutes after the last man had been through the reaction time test, vocational predictions were available for the entire group. Those who fell in the upper five or six tenths in the finishing score were immediately recommended for that work. Those who fell below this standard were given tests for handing out stock, if workers of that sort were needed at the time, otherwise they were given the test for general factory work. If they did well in this latter they were recommended for any general moderately skilled operation where there was need, but if they did poorly in this series, they were recommended for some unskilled laboring operation. If there was a need for clerical help, some of those who appeared more alert were given the clerical tests. This whole procedure would doubtless be different at a time when workers for some other operations besides finishing were badly needed. The general system with a large number of regression equations available would be to give first the tests for those occupations where there was the greatest need.

It occasionally happened that the recommendations on the basis of the tests ran counter to an individual's interest. A man naturally succeeds better at something in which he is interested. If, however, he preferred a certain type of work for which he was mentally or physically unfit, it was an economic loss both to himself and to the company to engage him for that work. When such a situation was found, effort was made to dissuade the man from the work which he wished to do. Failure in the work would doubtless cause a loss of interest, and the plodder who learns after a long time with economic loss to himself and to the company, is an exception. In many cases where the man appeared rather intelligent, the method of developing the tests and his probable expectation on the basis of his results were explained to him, so that he understood that his recommendation for some other occupation was due to the fact that he stood a poor chance of success in the one he preferred.

As the writer did not intend to remain with the factory permanently, a man was trained to carry on the work of testing applicants in the employment office. He first spent some time watching the writer give the tests and studied the instructions carefully. He then gave the tests himself under the supervision of the writer and all undesirable reactions were corrected. He was provided with a manual and tables for weighting the tests so that it was possible for him to obtain the weighted test scores with perfectly standardized method.

Effort was made during this month to follow up the men hired on the basis of test scores. Few of the men reached the piece-work stage during this period, but record was obtained of their daily production and in many cases estimates obtained from foremen as to their probable success. Had the writer remained on the ground longer, these estimates, of course, would have been much more systematized. Practically all the men who scored above the average in the tests and were hired on that basis were doing successful work when the writer left. A number of men with low test scores were hired as a check on the method. All of these gave up the work in the course of a short time, indicating that lack of success seemed to produce lack of interest.

The work was running smoothly when the writer left. Shortly after that time, there was a change in the executive personnel of the company,—the president and general technical superintendent, who had been the main sponsors of this work, being among those who resigned. At the present writing the work has been temporarily dropped, but it is hoped that it will later be resumed because at last accounts it was definitely vindicating itself.

10. *Conclusions*

The foregoing account of research at a rubber tire factory illustrates one method of psychological approach to the problems of vocational selection in industry. The general procedure was the standardization of mental and motor tests upon workmen of known vocational ability in order to use these standards in assigning applicants at the employment office to the type of work in which they stood the greatest chances of success. The operations at the factory seemed to fall into three large classes: those requiring no special mental ability, such as shoveling and trucking; those involving the learning of a few relatively simple operations and requiring a modicum of general intelligence; and those involving more definitely specialized mental or motor abilities. The last of these presented by far the largest and most difficult problem.

Effort was made in all cases to obtain the criterion, i. e., the vocational or production rating before tests were given at all. Ratings made by a given foreman were transformed into terms of the standard deviation of his ratings. They could then validly be averaged with estimates made by another foreman. A wide range of tests, both group and individual, were used. Certain of these were selected for each operation studied on the basis of tests of a preliminary sampling of good and poor workers of the sort under investigation. This selected set of tests was given to workers of all degrees of ability at a certain operation and each test score correlated with the vocational rating. The most promising tests were retained, the coefficients corrected for attenuation, partial correlations computed and the regression equation derived in order to weight the tests and get the best possible prediction of vocational ability. After determining the correlation of combined weighted test scores with the criterion a table was made with deciles of combined weighted test scores tabulated for deciles of probable vocational success. It was thus possible to predict from the test scores the probability of an applicant's being in the highest tenth, next highest tenth, etc. of workers in a given operation. These methods were applied to a number of rather specialized operations,—finishing, treading and building tires, handing out stock and clerical work. With the exception of building final correlations between tests and criterion of between .56 and .71 were obtained, thus giving considerable reliability to the predictions made on the basis of the tests. For the less specialized factory operations a general intelligence scale was devised to differentiate good from poor workers in the average moderately skilled operation and to further differentiate these workers from those fitted definitely for only unskilled labor.

These test scales were given to new employees and recommendations made on the basis of the decile distributions.

Effort was made to hire a person for a type of work in which he fell in the upper five or six tenths of a normal distribution of combined weighted test scores. If unsatisfactory in the tests for one operation he was tried in those for another until he *finally* made a good showing or else demonstrated his fitness for only unskilled labor. If recommendations ran counter to an individual's interests his chances of failure in the desired work were explained and an effort made to interest him in the recommended work. A man was trained to give the tests and continue the work after the writer left the factory. The progress in the factory of new help hired on the basis of the tests was, at last accounts, vindicating the methods.

WHAT INDUSTRY WANTS AND DOES NOT WANT FROM THE PSYCHOLOGIST[1]

By ELIOTT FROST, Rochester, N. Y.

Our topic is predicated upon the proposition that Psychology is both willing and eager to place at the disposal of industry—as it has already done for medicine, education, and the several emergencies of the War—its methods and its conclusions.

There was a time—sufficiently recent for these Emersonian walls to have heard and remembered—when to stress the practical application of our science, constituted a gaucherie frowned upon in some of the highest circles. If recent literature reflects inner conviction, that day is happily past, and consolation may be given the mourners by those of us for whom such a time never existed—for nothing good has really died. Our plight was due, first, to the youthfulness of our science, which felt it proper enough to be seen and not heard—but also to certain standards of scholarship professing to find in a liaison of theory with practice the liability of offspring for whom the assumption of paternity might prove an embarrassment. Medicine does not suffer as a science because men are healed thereby, nor need psychology regret finding itself similarly at the service of mankind. Such service as she can render will occasion fresh demands for service, and the incentive to search and research will grow with the years. It is because I assume, then and therefore, the response of the psychologist that I dare to speak for industry in terms so categorical.

Industry is worthy of such help as she needs. The needs of industry to-day are unprecedentedly great. These—I take it—are two self-evident propositions. A part of the dilemma of industry, however, is or shortly will become, the multiplicity of her doctors and the variety of their pottles. Let me cite for you a few examples from the row of industrial panacea. Each bottle is labeled by those who dispense it, as a "General Specific for Industrial Diseases."

You will not be misled by such a recital into the thought that these measures are worthless. On the contrary, I shall mention some which I believe to be important, not only, but indispensable. The list follows:

---

[1] Read at the Annual Meeting of the American Psychological Association, Cambridge, Mass., Dec. 31, 1919.

Americanization of the Alien
Establishment of Bonus Systems
Better Housing of Employees
Universal Continuation Schools
Increasing Efficiency Through Diet
Daylight Saving Laws
Efficiency Engineering in Factories
The Establishment of Foremen's Classes
Find Yourself Campaign for Working Boys
Graded Wage Scales
Compulsory Health Insurance
Industrial Democracy Systems
Special Legislative Enactments
Morale Work
Diagnostic Motion Pictures in the Shop
Nationalization of Basic Industries
National Board of Community Speakers
Psychological Tests
Questionnaires
Personnel Ratings
Systematic Propaganda
Recreational Athletics
Rest Periods and Fatigue Studies
Restriction of Immigration
Shop Committees
Employment Management
Socialism
Revision of the Metric System
Excess profits and other Taxation
Thrift Campaigns
Unionism
Vestibule Schools
Vocational Guidance
Welfare Work

This list could be doubled, nor is it a fictitious one. Each separate proposition is brought forward by earnest and honest men. Each one represents an actual proposal to the writer by some man, or group of men, within the last three months. Now, it is not the lack of excellence in these proposals, but the confusing number of them that constitutes a handicap. To be sure, many of them demand a leadership by unusual personalities or extensive and elaborate technique, or both. Even this is, however, the minor difficulty. The tendency to-day, despite the gravity of the problems faced, is to reject all panacea from a sheer inability to analyze and choose from among them. Industry is somewhat in the position of a man overboard, liable to perish from the bombardment of life-preservers thrown to rescue him.

Making money is, and broadly speaking must be, the expectation of business. The manufacturer must buy, manufacture and sell to show a profit or cease to do business. If the psychologist is to help, he must help make money. He must, in short, be both specific and practical.

Eight or ten years ago, industry was especially concerned with Workmen's Compensation Laws and with methods of Industrial Safety. Both these research problems are now solved with reasonable adequacy. On the other hand, the unrest following War has emphasized a new set of problems, the major of which are:

(a) Unionism, (b) Labor Turnover, (c) Selection and Training of Foremen, (d) Education of the Alien, (e) Wage and Hour Adjustments, (f) Housing, (g) Working Conditions, and in some States (h) Compulsory Health Insurance, (i) Taxation, (j) Continuation Schools.

Four of these have special psychological implications:

First, *Labor Turnover*—By Labor Turnover, we mean, roughly, the number of men hired in a given industry, per year, in proportion to the number of jobs. This varies all the way from zero to 8000%. 100% is a small, rather than a large, average. That means that the employer having, say a hundred workers during the twelve months, hires a hundred additional workers to fill vacated positions—some of these positions, of course, being emptied and filled many times over, while others are not disturbed. The annual economic loss in this country from Labor Turnover is estimated at one and one-half billion dollars, or $50 per head. Practically 50% of it is due to lay-offs or discharge. The discharge is usually for one of four reasons: Inadaptability, Unwillingness to Work, Wrong Attitudes Toward Work, or Positive Misconduct. The other 50% is due to dissatisfaction. Where turnover is high, dissatisfaction is the chief factor in the turnover. It is significant, on the other hand, that when the learning period in any job is long, turnover tends to be low; and, in general, of course, is lower in the skilled trades.

Psychology can help the employment manager in his hiring and firing and making of transfers within the plant, if it can devise a simple—preferably oral—readily applied test, to determine three things—(a) Intelligence, (b) Adaptability to Particular Tasks, and (c) Temperament.

In this connection, the well-disposed psychologist, however, must remember two hard cold facts: *First,* That labor in any community is usually either very abundant or abnormally scarce. If abundant, rough empirical tests, with recommendations do very well. If scarce, the man will be taken on anyhow whatever his mental rating. *Second,* We need to remember that vocational training does not *place* a man. The Law of Supply and Demand, and not his ambitions or aptitudes, determines the career of most men in industry.

A second timely problem, in which the psychologist has proper concern, is that of Americanization. This involves the education of the illiterate and of the foreign-born, both in English and in American ideals.

An analysis of racial psychology and the application of particular educational methods, devised to meet the educational problems of the several races, has not yet been made, so far as the speaker is aware. The failure of some of the most ambitious attempts to minister to industrial needs during the last four years, has been due to the failure properly to educate. At one plant, for instance, of 800 employees, in spite of an elaborate and particularly successful shop committee system, an analysis of the workers showed that 500 of the 800 were either foreigners or illiterates, and had no conception of the machinery set up in their behalf. They, therefore, went out on strike.

If the Swede, the *Pole,* the *Italian, the Greek,* are to be taught the advantages of industrial democracy, the lesson must be built upon the background of a proper racial psychology.

There is, for instance, more than a mild tendency to racial monopoly in certain industries. The French Canadians drift toward the cotton factories, copper mining, smelting, boots and shoes—the Croatians toward the mines, steel and the filthy trades—the Danes toward leather, furniture, collars and cuffs—the Dutch toward silk-making, dyeing, furniture factories, and the like—the Armenians toward cigarette making and peddling—Greeks into the blacksmithing, baking, shoemaking, boot-blacking and the like—Hebrews into the small manufacturing of the sweat-shop variety—Chinese into importing—French-Swiss into the hotel and restaurant business, silk industries and embroideries, and so on.

Again, certain races are easily assimilated. For instance, the French-Swiss, Ukrainians, Russians, Poles and Japanese. Others are distinctly clannish, as the Slovacs, Armenians, Albanians, Portugese and Chinese. Such social characteristics of the various nationalities are clearly reflected, both in their work, and in their relation to the industry.

The laborer's attitude toward industrial relations is determined by his nationality more than any other single factor. The Jew, for instance, demands an arrangement in which he can bargain. He is continually thinking of how much he is receiving for his labor. He is really more conscious of his labor organization and its methods than he is of his creative labor faculties. As a thinker, he is usually of the radical stamp. The Italian's highly emotional nature lends itself readily to directions by the organizers. It is the testimony of the executives that he cannot be trusted without reservations, and that he is apt to be sullen and moody. The German workman is of placid disposition, loves detail, is particularly effective on precision work. The Pole and Croat usually do the dirty work in the plant. For instance, in the grinding of lenses, a job in which the workman becomes covered with damp red pumice, the Poles predominate.

In any Americanization program then, educational methods must vary with nationalities involved. The ideal method of reaching the slow-moving, generally placid Swede will not serve for the emotional Italian; nor that adapted to the keen-witted Jew for the average German mechanic in the shop.

These facts and tendencies indicate a distinct problem in race psychology, but one applicable in its solution to the workman at the bench. Racial characteristics affect quality and quantity of work. They affect adaptability, willingness to work, attitude toward work, and misconduct—the four basis criteria of discharge.

In short, an industrial psychology has still to be written, although the need for it has long been great.

In connection with the Continuation or Part-Time School, psychology, and especially educational psychology, can render a third service to industry. Such a school, inaugurated some 18 years ago in Chicago, is now permissive in Massachusetts, Ohio, Michigan and several other States, and compulsory in Wisconsin, Pennsylvania and New York. It provides for the continued schooling, on the employer's time, either in the arts and sciences, in civics, or in vocational work, from 4 to 8 hours a week, of all youth in industry, and who have no education beyond the grade school.

If this Law is not to operate as a handicap upon the employment of such workers, it must justify itself by the increase of efficiency through a sound educational program. This efficiency will be measured by the employer—whether we like to face this fact or not—in terms of dollars and cents. Can, then, the psychologist offer a suggestion as to the best method of instruction, the best time of day within the factory schedule, the wise allocation of hours during the week? Can he advise as to whether such instruction should be by racial groups, or to a mixed group of all races? Can he hope to show the employer a relation between broad education and production units?

The fourth and perhaps most important service renderable to industry by the psychologist, concerns the foreman. The War Labor Board claims that 98% of industrial disputes faced during the War, led back to the foreman. The foreman is the Army Top Sergeant. The passed buck reaches him. He must handle men mentally and physically. Until recently, the emphasis was upon the physical. Under newer ideas and ideals of persuasion, therefore, industry now finds itself saddled with men, whom it has taught to be " hard-boiled." These petty officers, either in industry or War, can make or break morale. Foremen, therefore, and prospective foremen, must be selected and trained for their jobs. The attitude between the foreman and the man at the bench, depends in part, upon the power the foreman holds. If he can hire and fire at will, methods of persuasion are less likely to be used. Two modern industrial tendencies are helping in the solution of this problem:

First, The centralization of employment, whereby the hiring and firing is taken out of the foreman's hands; and

Second, The vestibule-school idea wherein the foreman becomes a teacher to the man at the bench.

These are both steps in the right direction. While the major task of the true foreman is that of problem-solving, he will always remain a manipulator of human material, and the product which leaves his room will continue to reflect his success or failure in dealing with psychological factors. At present, the average foreman is rarely a teacher, but he is a man, usually, who can DO. Tests devised by the psychologist to aid in the selection of foremen should, then, have in mind three qualifications:

First, Ability to solve problems.

Second, Ability to handle men.

Third, Ability to teach the theory of that which one knows in practice.

Industry does not want a booklet of psychological tests. The advertised success of the latter in the Army Camps has not sold the idea to the manufacturer. He must be shown the value of any technique by patient education and proved results. The psychologist must not become "cocky." He is already accused of it. His proposed solutions must check with practical experience. He must know concrete conditions; must devise his remedial laboratory experiments in the light of them, and not *sui generis*. No time ever needed theories less than ours, and no time ever needed facts more.

If we are to remove the dust from Industrialism and give the day's work its ancient sparkle, the modern tendency toward de-personalization and mechanization must be checked. When this tendency succeeds, incentive fails—the adventure of creative work departs—business captures industry for profit—ideas become institutionalized, and the workman a machine.

More than tests, analyses, teaching of aliens, or training of foremen, the psychologist is helping industry when, knowing its problems, he brings to their solution a new vision of the importance of psychology itself.

A CONSTANT ERROR IN PSYCHOLOGICAL RATINGS

By Edward L. Thorndike, Teachers College, Columbia University

In a study made in 1915 of employees of two large industrial corporations, it appeared that the estimates of the same man in a number of different traits such as intelligence, industry, technical skill, reliability, etc., etc., were very highly correlated and very evenly correlated. It consequently appeared probable that those giving the ratings were unable to analyze out these different aspects of the person's nature and achievement and rate each in independence of the others. Their ratings were apparently affected by a marked tendency to think of the person in general as rather good or rather inferior and to color the judgments of the qualities by this general feeling. This same constant error toward suffusing ratings of special features with a halo belonging to the individual as a whole appeared in the ratings of officers made by their superiors in the army.

The official rating plan devised by Walter Dill Scott called for separate ratings for Physical Qualities, Intelligence, Leadership and Personal Qualities (i. e. Character). The instructions very emphatically required each of these four to be estimated independently of the others, as appears from the directions quoted below. Yet the correlations of the Intelligence rating with the ratings for Physique, Leadership and Character made by a very conscientious officer in the case of 137 aviation cadets whose work he, as flight commander, supervised, were .51, .58 and .64 respectively. These are all higher than reality, plus the attenuation due to erroneous judgments, could well give, especially within the restricted range of the commissioned-officer group. They are also too much alike. In reality Intelligence and Character or Intelligence and Leadership should give about three times as close a correlation as Intelligence and Physique.

> " How to Make the Scale.
>
> 3. Make a list of about a dozen officers of your own rank and not above the average age of officers of this

rank. They should be men with whom you have served or with whom you are well acquainted. Include officers whose qualifications are poor or mediocre as well as those who are highly efficient. This list serves merely as a convenient reservoir of names; the names actually used in the scale may include others.

4. Look over your list from the viewpoint of *Physical Qualities* only. Disregard every characteristic of each officer except the way in which he impresses his men by his physique, bearing, neatness, voice, energy and endurance. Select that officer who surpasses all the others in this qualification and enter his name on the line marked highest under Physical Qualities. Now select the one who most conspicuously lacks these qualities and enter his name on the line marked lowest. Select the officer who seems about half way between the two previously selected and who represents about the general average in physical qualities; enter his name on the line marked middle. Select the officer who is half way between the middle and the highest; enter his name on the line marked high. Select the one who ranks half way between middle and lowest; enter his name on the line marked low.

5. In the same manner make out scales for each of the other four sections (Intelligence, Leadership, Personal Qualities and General Value to the Service).

How to Use the Scale.

6. Rate your subordinate for Physical Qualities first. Consider how he impresses his men by his physique, bearing, neatness, voice, energy and endurance. Compare him with each of the five officers in section I of the Rating Scale, and give him the number of points following the name of the officer he most nearly equals. If he falls between two officers in the Scale give him a number accordingly (e. g. if between Low and Middle give him 7, 7½ or 8).

7. Rate the subordinate in a corresponding manner for each of the other four essential qualifications. Under III (Leadership) and V (General Value to the Service) consider which officer he will most nearly equal after equivalent experience.

* * * * * * * * *

Points for Special Attention.

9. In making or using any section of the scale, consider only the qualification it covers, totally disregarding all the others.

> I. Physical Qualities.
>
> Physique, bearing, neatness, voice, energy and endurance. (Consider how he impresses his men in the above respects.)
>
> II. Intelligence.
>
> Accuracy, ease in learning, ability to grasp quickly the point of view of commanding officer, to issue clear and intelligent orders, to estimate a new situation, and to arrive at a sensible decision in a crisis.
>
> III. Leadership.
>
> Initiative, force, self reliance, decisiveness, tact, ability to inspire men and to command their obedience, loyalty and co-operation.
>
> IV. Personal Qualities.
>
> Industry, dependability, loyalty, readiness to shoulder responsibility for his own acts, freedom from conceit and selfishness, readiness and ability to co-operate.
>
> V. General Value to the Service.
>
> His professional knowledge, skill and experience; success as an administrator and instructor; ability to get results."

The same effect appears in the ratings given by other officers. The correlations are too high and too even. For example, for the three raters next studied the average correlation for physique with intelligence is .31; for physique with leadership, .39; and for physique with character, .28.

The same constant error appears in the correlation of the total Scott rating with a rating for technical ability as a flyer in the case of aviation officers. It is known from abundant evidence that technical ability as a flyer is a rather highly specialized quality.[1] Considering the restricted range of the aviation cadets, the correlation between general ability for officer work and technical ability as a flyer could hardly be above .40, without any attenuation. As attenuated by the imperfections of the rater's knowledge of both, it could hardly be above .25. Yet the correlations for the eight raters studied in this respect are .74, .85, .52, .91, .63, .72, .47 and .53, an average of .67. Obviously a halo of general merit is extended to influence the rating for the special ability, or vice versa.

Mr. Knight of Teachers College has studied this same effect in the case of 129 teachers rated by their superior officer for certain qualities on the Boyce score card. The ratings in question were official and were used to determine salaries and promotions. General merit as a teacher has correlations of .68 with intellect, .79 with power in discipline, and .63 with voice. It is clear that the rating of a teacher's voice must have been influenced by the general impression of her ability. Voice correlates .50 with "Interest in Community Affairs," and .63 with intelligence!

---

[1] See "The Selection of Military Aviators: Mental and Moral Qualities," *U. S. Air Service*, June 1919.

The correlations reported in the original study by Boyce show this same effect. General merit as a teacher is reported to correlate as follows:

With	General Appearance	47
"	Health	56
"	Voice	53
"	Intellect	62
"	Initiative and self reliance	77
"	Adaptability and resourcefulness	80
"	Accuracy	74
"	Industry	69
"	Enthusiasm	71
"	Integrity and sincerity	63
"	Self control	66
"	Promptness	66
"	Tact	69
"	Sense of Justice	61
"	Academic preparation	41
"	Professional preparation	38
	etc. etc.	

(The last is the lowest of the forty-five correlations reported.)

In the cases so far the correlations are a resultant of (1) the real facts, (2) the constant error of the "halo," as we may call it, and (3) the reverse error of attenuation due to chance inaccuracies in the ratings. In certain further work by Mr. Knight the correlations are freed from the last influence, by being based on the composite rating of two groups, each of a number of teachers who knew the individuals to be rated fairly well. The self-correlations of the ratings by one such a group with ratings for the same trait by the other group are over .90. The correlations for general ability as a teacher with intellect and with ability to discipline are about .95 and .80! The correlation of intelligence and ability to discipline is about .80! The correlations of a standard test of intelligence with general ability as a teacher and with ability to discipline are, for the individuals in question, not over .3.

The writer has become convinced that even a very capable foreman, employer, teacher, or department head is unable to treat an individual as a compound of separate qualities and to assign a magnitude to each of these in independence of the others. The magnitude of the constant error of the halo, as we have called it, also seems surprisingly large, though we lack objective criteria by which to determine its exact size. As a consequence science seems to demand that, in all work on ratings for qualities the observer should report the evidence, not a rating, and the rating should be given on the evidence to each quality separately without knowledge of the evidence concerning any other quality in the same individual.

PSYCHOLOGICAL TESTS AS DIAGNOSTIC OF VOCATIONAL APTITUDES IN COLLEGE WOMEN

By ELSIE MURRAY, Sweet Briar College, Virginia

The use of tests in the diagnosis of the academic difficulties of college freshmen, or in the sifting of the qualifications of applicants for admission, is now sufficiently familiar. Their application to the differential study of members of the graduating classes and of their fitness for various fields of activity is less widespread. Under the stress of war-time conditions, and the resultant necessity of recommending women graduates to a wide variety of positions, the following attempt to extend and adapt the technique of psychological testing to the latter problem was made by the writer.

Tests Utilized. Four Series. A series of tests already under way in the junior and senior classes of a woman's college (numbering 36 and 32 respectively), in illustration principally of certain routine topics in the courses in general, experimental, social, and educational psychology, was further extended. The tests were then divided into four sets according as they were judged roughly from past experience to be suitable measures of the ability to handle ideas, symbols, things or people (captions borrowed with certain alterations from Thorndike). Each set, after considerable shifting, contained (with the exception of the fourth) ten distinct tests, representing roughly two hours actual classroom time each, distributed over a period of about a year and a half. The tests utilized were for the most part standard ones, supplemented by a few original with the author. The group method was usually employed, though a few were individually administered. The first series contained tests in *controlled association, logical memory, completion, proverb matching, generalisation, information,* etc.; and will be referred to as the Reasoning or R series. The second was made up of *rote memory, rote learning, addition, spelling, cancellation, easy directions,* and related tests, and will be known as the Accuracy or A series. The third, or "Practical Ability" or P series, contained tests of motor dexterity and alertness to surroundings, such as *tapping, card sorting, aussage, construction puzzle, weight suggestion, divided attention,* etc. The fourth, or "Social Ability" set, comprises a series involving judgments upon *personal attributes, situations,* and *facial expression of the emotions,* referred to as S series.

Evaluation of the Scores. Pooling. The results were evaluated as follows: Each test was scored on a point system (empirically determined and somewhat approximating a percentile ranking) for both speed and accuracy, the maximal credits assigned to any test being usually ten points. The scores within each set were then pooled, and each student assigned a rank (1 to 36 or 1 to 32) on the basis of each of the four combined scores (R, A, P and S). While a small number rank high on all four ratings, this procedure was successful in differentiating each class roughly into four groups: one, of those doing better in R than in any other series; another, in A; another in P, etc. The correlations[1] between

$$\rho = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

the scores for the four test series range from +.05 for A and S to +.68 for R and S.

Interpretation of the Resultant Ratings. The significance of the differentiation of individuals thus secured was measured roughly by comparison of test ratings with classmates', instructor's, and self estimates of I (intelligence), A, P and S abilities, and with academic grades. The classmates' estimates were obtained by having each subject arrange the names of all the members of her class in order (1 to 36), according to the degree of the ability in question which she believed each to possess. The median position of each individual for each trait was then calculated, and a composite ranking representing the group judgment obtained for *general intelligence, accuracy, practical* and *social ability,* respectively, for both classes; an additional estimate on *teaching ability* for 1918. This method (of relative position) was found to yield higher reliability coefficients (ranging from .82 for S to .97 for I) than were secured by the seven or nine grade system.

1. The correlations of *test rankings* with the *corresponding estimates* were as follows, for the R, A, P and S abilities: for 1918, .63, .30, .12, and .19 respectively; for 1919 (with im-

---

[1] All correlations were figured by the Spearman 'Rank-Differences' formula.

proved methods of giving and scoring) .73, .36, .53, .26. The corresponding correlations for tests and instructor's estimates (the writer's was the only one available)[2] are, for 1918, .84, .79, .50, .44; for 1919, .83, .71, .59, and .65. While the coefficient for reasoning tests and intelligence estimates (both classmates' and instructor's) is high (averaging .68 and .835 for the two classes),[3] as compared with similar measures quoted by others (see Hollingworth, for classmates, .62; Kitson, Ruml, Thurstone, Abelson, and Webb, for instructors, average .57, .66, .60, .58, and .545, respectively), the correlations for the A, P and S abilities are less striking. The latter are, however, for the most part higher with the improved methods of 1919; higher also for instructor's than for classmates' estimates.

The real significance of these figures (slight, it must be admitted, in certain cases), the possible weaknesses of both tests and estimates, as well as certain differences between faculty and student estimates, are best brought out by the complete tables of correlations and cross-correlations. Certain inherent shortcomings of either P tests or estimates, for instance, would seem to be indicated by the fact that the P estimates of classmates and instructor correlate as well or somewhat better with A than with P test scores. On the other hand, the closer correlation of classmates' (not instructor's) A estimates with P than with A test scores is probably to be attributed merely to a slip in the instructions forming the basis of the A estimates in *1919*, in which manual rather than clerical accuracy was emphasized. Finally, the higher correlation of classmates' estimates of S ability with R than with S test scores challenges explanation: while undoubtedly due in part to an inadequate differentiation of the R and S test series, it is in part also to be attributed to the difficulty of securing reliable estimates of the quality aimed at, the students tending in many cases to substitute a judgment based upon superficial amiability of manner, for the actual estimate of tact and sympathetic understanding of human nature which was sought.

---

[2] The correlation coefficients between instructor's and classmates' estimates for I, A, P and S average .77, .61, .28, and .35 for 1918 and 1919.

[3] The exact degree of correlation involved may be clarified by the following statement: the coefficients of .63 and .73 for class I estimates and R tests mean respectively that, in 1918 and 1919, 72 and 81% of those scoring above the median in the tests, ranked also above the median in the class estimates.

CORRELATION TABLES FOR CLASS OF 1919

	CLASSMATES' ESTIMATES				INSTRUCTOR'S ESTIMATES				
Tests	I	A	P	S	I	A	P	S	T
R	_.73_*	.19	.52	.42	_.83_	.51	.49	.57	.80
A	.45	_.36_	.56	.05	.42	_.71_	.74	.04	
P	.53	.43	_.53_	.32	.55	.58	_.59_	.38	
S	.39	−.23	.14	_.26_	.66	.18	.26	_.65_	

* With the one exception indicated, these are all positive.

2. The correlation of *test ratings* and *academic standing* was next attempted, bringing to light an apparent variation in the reliability of grades from year to year, or class to class; further, the relative unreliability of freshman as compared with upper class marks. The correlation, for instance, of academic standing with R tests drops from .71 in 1918 to .46 in 1919; with instructor's estimate of intelligence, from .80 to .68; with classmates' from .86 to .58; with instructor's estimate of teaching ability, from .81 to .55 (and this at the same time that the coefficients for classmates' estimates of intelligence and R tests rise from .63 in 1918 to .73 in 1919; for instructor's estimate of teaching ability with R tests from .70 to .80, with combined R A P S scores from .75 to .82).

This collapse of the correlation coefficients for academic standing in 1919 was traced to two sources. First, to the fact that for this class freshman and sophomore grades only (in required courses) were utilized, while in 1918 junior grades had been included; elimination of the latter in calculating 1918's standing reduces the correlation with R tests from .71 to .60, with intelligence estimates (classmates'), from .86 to .77. More particularly, the drop in the 1919 coefficients appears to be correlated with the presence in the class of two members of considerable natural ability, but of indifferent academic ambitions, especially in the earlier years of their college course; elimination of whose scores raises all of the 1919 grade correlations. In the opinion of the writer, the facts warrant the assumption of the superior value, year in and year out, of R test ratings over academic grades, as measures at least of potential ability.

Corroboration of our hypothesis as to the superior reliability of grades from the later college years is apparently furnished by the corresponding correlations (of R tests and grades) reported by others workers; e. g., .20 for freshmen, .44 for a mixed group of freshmen, juniors, and seniors (Kitson); .42 for juniors, .57 for seniors (Hollingworth)—note the ascending series. The possible explanation, that mechanical memory work is the determining factor in freshman and sophomore grades, as against initiative and reasoning ability in the upper classes, is not, however, supported by the relatively low correlation of grades and A test scores for 1919.[4] Freshman instability and irresponsibility, or variations in the standards and personnel of the instructing force are presumably the principal sources of the discrepancies and inconstancies noted.

Scrutiny of the table of correlations for *test scores* and *grades in particular subjects* brings to light other facts of interest. For 1919 (the test scores for which are regarded as more reliable), the study correlating most closely with R test scores is English (+.74); this subject, moreover, gives decidedly lower correlations (+.31 to .36) with all other test sets. The study yielding the highest correlation with A test scores is Mathematics; the figure, to be sure, is low (+.34) and is followed closely by the Latin and English coefficients. However, the correlation of Mathematics with all other test groups is lower (+.02 to —.19). While the subject correlating most nearly with the P tests is English (+.34), Science is a close second (+.31, P. E. =.105); moreover, the correlation of Science with all the other test sets is lower (+.26 to —.01). Lastly, the subject correlating the most closely with S scores is History (+.64), and the correlations of the three other sets of scores with this subject are all lower (.57 to .10).

CORRELATION OF TEST SCORES AND COLLEGE SCORES

Tests	Latin	Mathematics	Science	English	History	Modern Language
R 1918	.67	.29	.39	.70	.60	
1919	.31	.01	.21	*.74*	.57	.37
A 1918	.59	.39	.57	.63	.41	
1919	.32	*.35*	.26	.31	.23	.20
P 1918	.10	.13	.24	.27	.26	
1919	.21	.02	.31	*.34*	.10	.10
S 1918	.45	.0	.18	.57	.47	
1919	.23	-.19	-.01	.36	*.64*	.30

Whether or not these correlations will be found valid for other colleges or classes remains to be proven. The relationships anticipated had been, it may be noted, between Latin and R tests, Mathematics and A tests, Science and P tests, and either English or History and S tests. The significance of the (actual) English and R test correlation is ambiguous. Does it indicate a combination of R tests too closely dependent upon verbal facility and vocabulary?[5] the superior *value of achievement in freshman English* as a criterion of *general ability?*; or *merely* the superior *reliability* of the English instructors' judgments? The virtually zero correlation of Mathematics grades (for 1919) with all but A tests gives rise to further speculation. Is it possible that success in college Mathematics presupposes, at least in women, a specialized talent existing, as often as not,[6] at the expense of or in inverse ratio to humanistic interests and practical abilities?

---

[4] The complete correlation table for test scores and grades is as follows: for 1918, for R, A, P, S and total test scores, .71, .64, .31, .48, and .75; for 1919, .46, .33, .26, .24, and .45.

3. Finally, each student was asked to register the line of work for which she believed herself to be best adapted. In 1918, when an option of three lines of work only was offered, in a class of thirty-six, 11 registered for teaching, 13 for clerical or secretarial work, 12 for practical. Of the 13 registering for clerical work, 8 (61.5%) score in the upper third of the A test ranking; of the 11 assigned to teaching, 6 (54.5%) appear in the upper third of the R test ratings; while of the 12 signing for practical work, only 3 (25%) are found in the upper third of the P test ranking, although 48% score higher in this than in any other test rating. In 1919, of those mentioning clerical work, 62.5% rank in the upper third of the A tests; of those mentioning teaching, 50% in the upper third of the R tests; of those mentioning social work, 46% in the upper third of the S test scores; of those signing for practical work, 37.5% in the upper third of the P test rating, while 43.5% do better in P tests than in any other set.

Conclusions. From the point of view of vocational guidance, the present study can be regarded as tentative only; both the experimental procedure and the securing of estimates demand modification, and checking up by post-graduate study of the individual. At the same time, the probable existence of distinct aptitudes which fit or unfit the individual for one of the four lines of work mentioned seems to be fairly demonstrated by the degrees of difference manifest in the four test ratings. Whether these aptitudes are innate or acquired, whether they depend upon specialised abilities (unit traits), upon interests derived from a variety of factors, or are merely the by-products of varying degrees of general ability, is not our immediate concern. Careful weighing of the data in hand, however, inclines the writer to the view that degrees of intelligence, seconded by certain physical and temperamental peculiarities, play an important rôle. While certain gifted individuals may display ability along all four lines, only those lacking in ideas and originality are likely to satisfy themselves for any period with purely clerical activities; few possessing them are likely to content themselves with, or give undivided attention to, work largely manual or practical. The actual bar to clerical or practical activities, however, may be primarily defective eyes, a nervous temperament, or natural clumsiness of hand. It is, moreover, an interesting speculation as to whether variations in instinctive endowment, excess or deficiency in the social and sympathetic instincts, in curiosity (as manifest in alertness of the senses, and a natural turn for manipulation), in facile sensory motor learning, may not play a part in determining inclination toward or away from social, practical, or intellectual work; or whether a differentiation in the degree of development of the motor centres for the larger and the smaller muscles may not lie at the root of a preference for or an aversion to practical or clerical work.

---

[5] Against this construction may be cited the high correlations of English grades with instructor's I estimates (.71 and .64 for 1918 and 1919); and with classmates (.70 and .77).

[6] In this connection, it is interesting to note that the correlation between freshman Mathematics and English for 1918, is +.555, for 1919, only +.08. It is possible, of course, that in larger, more representative groups this relationship would exhibit more constancy. The corresponding figure quoted by Tolman (*Journal of Educational Psychology*, February, 1919) for 27 freshmen is .47±10; by Uhl, for 100, .48, *ibid.*, Jan. 1919; by Bell *ibid.*, Sept. 1916, for 750, .46.

Pending more definite knowledge along these lines, the use of class estimates and of tests of the variety here described will be found invaluable in stimulating the student to more careful weighing of alternatives and qualifications in the choice of a vocation; in accentuating the need of more adequate preparation; and, incidentally, in demonstrating the value of a wider range of choice in electives, to counteract the tendency toward too narrow specialisation so often revealed.

The time and labor thus expended may be further justified by the amount of light thrown upon undergraduate problems of a more purely academic nature by the test data. General tendencies, such as too complete a reliance upon verbal memory, defective habits of concentration or self-control, lack of expressive facility, of abstract or general ideas or of the ability to form spontaneous generalisations, habits of carelessness, the tendency to sacrifice accuracy to speed, an undeveloped social consciousness, inefficiency in handling one's ordinary working materials, may all be isolated and magnified by the range of tests here advocated, and the student advised accordingly. If the individual is looking forward to teaching, or to clerical or secretarial or social work, the professional motive may be utilized to stimulate effort along lines where weakness in the necessary elementary processes or qualifications can be demonstrated.

In the case of students involved in academic difficulties, the R tests may disclose a coefficient of intelligence so low as to render a satisfactory solution of their college problems unlikely; at the same time their A, P or S scores may indicate a practical (or other) ability sufficient to warrant the attempt to divert their ambitions from the conventional teaching career to lines of work more suited to their actual abilities. On the other hand, the discovery of a high intelligence coefficient in the unambitious may be utilized to induce the adoption of a more serious attitude toward college opportunities.

Summary

1. Scores obtained from a series of tests distributed over a number of months in the junior and senior years, when pooled in four sets to measure General Intelligence (or Teaching Ability), Accuracy (or Clerical Ability), Practical and Social Ability, respectively, afford evidence of decided dissimilarity in the four functions measured.

2. Classmates' estimates of each other's ability furnish additional information of vocational interest, and to a certain extent corroborate the findings of the tests.

3. College grades based upon freshman and sophomore courses appear on the whole to be less reliable than those of later years[7] as indicators of real ability and differentiation of talents.

4. Comparison of the student's own vocational choice with the various test scores reveals a fairly high degree of correspondence between individual ambition and experimental findings.

5. The test data accumulated are applicable to a variety of academic problems.

BIBLIOGRAPHY

1. Bell, J. C.: Mental Tests and College Freshmen, *Jour. Educ. Psych.*, VII, 1916.
2. Bingham, W. V.: Some Norms of Dartmouth Freshmen, *Jour. Educ. Psych.*, VII, 1916.

---

[7] This point is developed further, with reference to the major subjects pursued by the individual, in a fuller account of this investigation to be published later.

3. Bingham, W. V.: Mentality Testing of College Students, *Jour. Applied Psych.*, I, 1917.
4. Brown, W.: Some Experimental Results in the Correlation of Mental Abilities, *Brit. Jour. Psych.*, 1910.
5. Cattell, J. McK.: Homo Scientificus Americanus, *Science*, XVII, 1903.
6. Cattell, J. McK. and Farrand: Physical and Mental Measurements of Students, *Psych. Rev.*, III, 1896.
7. Downey, J.: Form and Position in Handwriting, *Jour. Educ. Psych.*, 1915.
8. Folsom, J. K.: Statistical Study of Character, *Ped. Sem.*, XXIV, 1917.
9. Healy, W., and Fernald, C.: Test for Practical Mental Classification, *Psych. Rev. Mon.*, XIII, No. 54, 1911.
10. Hollingworth, H. L.: *Vocational Psychology*, 1916.
11. Kitson, H. D.: Scientific Study of the College Student, *Psych. Rev. Mon.*, XXIII, 1, 1917.
12. Link, H. C.: An Experiment in Employment Psychology, *Psych. Rev.*, XXV, 1918.
13. McCrory, J., and King, I.: Freshman Tests at the University of Iowa, *Jour. Educ. Psych.*, IX, 1918.
14. Norsworthy, N.: On the Validity of Judgments of Character, *Essays in Honor of William James*, 1908.
15. Pintner, R.: Standardization of Knox Cube Test, *Psych. Rev.*, XXII, 1915.
16. Rowland, E., and Lowden, G.: Psychological Tests at Reed College, *Jour. Exp. Psych.*, 1916.
17. Ruml, B.: Reliability of Mental Tests in the Division of an Academic Group, *Psych. Rev. Mon.*, XXIV, 1917.
18. Sharp, S.: Individual Psychology, *Amer. Jour. Psych.*, X, 1899.
19. Simpson, B. R.: Correlations of Mental Abilities, 1912.
20. Stenquist, Thorndike, and Trabue: Intellectual Status of Children who are Public Charges, 1915.
21. Stern, W.: *Differentielle Psychologie*, 1911.
22. Terman, L. M.: *Measurement of Intelligence*, 1916.
23. Thorndike, E. L.: *Introduction to the Theory of Mental and Social Measurements*, 1913.
24. Thorndike, E. L.: Educational Diagnosis, *Science*, 1913, p. 133-.
25. Thompson, H. B.: *Mental Traits of Sex*, 1903.
26. Washburn, M. F.: A Study of Freshmen, *Amer. Jour. Psych.*, 1917.
27. Webb, E.: *Character and Intelligence*, 1915.
28. Weidensall, J.: *The Mentality of Criminal Woman*, 1916.
29. Whipple, G. M.: *Manual of Mental and Physical Tests*, 1915.
30. White, W. A.: *Outlines of Psychiatry*, 1912.
31. Whitley, M. T.: *An Empirical Study of Certain Tests for Individual Differences*, 1911.
32. Wissler, C.: Correlation of Mental and Physical Tests, *Psych. Rev. Mon.*, III, No. 16, 1901.
33. Woodworth, R. S., and Wells, F. L.: Association Tests, *Psych. Rev. Mon.*, XIII, 57, 1911.
34. Woolley, H. T., and Fischer, C. R.: Mental and Physical Measurements of Working Children, *Psych. Rev. Mon.*, XVIII, 77, 1914.
35. Yerkes, R. M., Bridges, J. W., and Hardwick, R. S.: *A Point Scale for Measuring Mental Ability*, 1915.

THE APPLICABILITY OF MENTAL TESTS TO PERSONS OVER FIFTY YEARS OF AGE

By JOSEPHINE CURTIS FOSTER,[1] *Formerly Chief Psychologist,* and GRACE A. TAYLOR,[1] *Formerly Psychological Interne,* Psychopathic Hospital, Boston

In psychological examinations made at the Psychopathic Hospital it has for some time seemed evident to the examiners that the majority of patients over 50 years of age do especially poorly in the tests of memory and in certain allied tests. Casual observation and another study[2] showed further that this deficiency depends relatively little upon type of mental disease and relatively much upon chronological age. The present paper is an account of our attempt to find reasons and numerical expression for the differences among our patients. We have attempted also to check our results by studying the influence of chronological age in the case of normal persons.

Cases Used and Their Comparability

The following groups of cases have been studied. No foreign-speaking adults and no children of foreign-speaking parents are included in any group.

(*a*) 106 older[3] patients in a large general hospital in Boston (not an insane hospital).

76 per cent of this group are men. So far as our problem

---

[1] Our thanks are due to the following physicians whose cooperation made possible the work on the normal old persons: Drs. John J. Dowling, Edmund W. Wilson, John L. Ames, John Bapst Blake, W. E. Faulkner, E. N. Libby, Edwin A. Locke, H. A. Lothrop, Paul Thorndike, and Townsend Thorndike. We wish also to thank Dr. F. H. Thomas for allowing us to work at the Foxboro State Hospital, and Dr. R. M. Yerkes for allowing us to use the data on school children.

[2] To be published shortly by J. C. Foster under the title "Significant Responses in Certain Memory Tests."

[3] Older, in this paper is taken to mean 50 or more years of age. Younger is to be understood similarly as less than 50 years of age.

is concerned, however, a study of the cases reveals no significant influence of sex upon results.

This hospital charges up to $10 a week where payment is exacted, but admits patients free when their families are unable to pay for treatment. 59 per cent were manual laborers. The average social status of the patients of this group is, therefore, relatively low. 32 per cent of them, for instance, left school before the 5th grade, and 85 per cent of them before the 9th grade.

All of these patients were given the Yerkes-Bridges Point Scale by the second writer. Our original plan was to include also several additional memory tests, but they had to be abandoned. In the first place, many of the patients had too poor eyesight to be given any "visual verbal" tests. But more than this, many of the older persons objected to attempting any memory test. At the mere hearing of the directions "when I am through reading I want you to tell me as much as you can of what I read" many of this group made such remarks as "Oh, no! I couldn't do that" or "Oh, no! my memory is too poor for that" or "That is child's play, don't try anything like that on me." The true reasons for refusal evidently were the conviction that their memory was poor and a reluctance to display this weakness. In general we found in working with normal older persons a similar reluctance to undertake other "difficult" tests. Some half dozen cases, for example, had to be discarded because the patient refused to attempt any test after the twelfth, that is to say, when the thirteenth was given, they would not attempt it, nor would they attempt any succeeding test. We were therefore forced to be content with giving the Point Scale alone.

(*b*) 315 normal young men between ages 20 and 30. This group is distinctly inferior in social status and education to the above group. 79 per cent were manual laborers. 65 per cent of them left school before reaching the 5th grade, and 96 per cent left before reaching the 9th grade. All of these cases were given the Point Scale (omitting tests 14 and 18) by other examiners, but all scoring was checked by us. Tests 14 and 18 were omitted because the majority of the group had too little schooling to be able to do any of the tests involving sentence construction.

(*c*) 316 normal school children between ages 10 and 19. These cases were from the Cambridge public schools and were tested in obtaining the original norms for the Point Scale.[1] Their average social status is at least as high as that of group (*a*).

(*d*) 136 older persons, patients diagnosed psychotic either at the Psychopathic Hospital or at the Foxboro State Hospital. Their social status is comparable with the patients of group (*a*). It is possibly slightly superior. 60 per cent of the group are men. 27 per cent only were manual laborers. 32 per cent left school before reaching the 5th grade and 75 per cent before reaching the 9th. All were given the Point Scale by examiners of the Psychopathic Hospital.

(*e*) 151 *younger* persons, patients diagnosed psychotic at the Psychopathic Hospital. 52 per cent of the group are men. 22 per cent only were manual laborers. 17 per cent left school before reaching the 5th grade and 74 per cent left before reaching the 9th. There is no reason to suppose that the social status of this group is different from that of the above group (*d*). All were given the Point Scale by examiners at the Psychopathic Hospital.

Comparison of Older Psychotic and Younger Psychotic Persons (Groups d and e)

If we temporarily disregard type of mental disease, we find that the younger group attains on the average a higher score on each test as well as a higher total score. The superiority of the younger patients is, to be sure, most marked in tests 13 (words in three minutes), 14 (three words in one sentence), 16 (drawings from memory), and 18 (dissected sentences), but we cannot be sure how much the results are affected by type of disease and by total score (or mental age).

If, in an attempt to eliminate the influence of total score, we fractionate our cases by total score groups, we find two such score-groups sufficiently large for use,—53 to 60 (mental age 10) and 72 to 79 (mental age 13). The data for these groups are given in Table 1.

---

[1] See Yerkes, Bridges, and Hardwick, *A Point Scale for Measuring Mental Ability*, 1915, Chap. 4.

TABLE 1

AVERAGE SCORES OF OLDER AND YOUNGER PSYCHOTIC PERSONS

Total Score		53–60		72–79	
Age Groups		Older	Younger	Older	Younger
Number of Cases		15	18	7	21
Test	1	3.0	3.0	3.0	3.0
	2	3.1	3.2	3.7	4.0
	3	3.0	2.9	2.8	3.0
	4	3.2	3.2	4.1	4.0
	5	2.8	3.9	4.0	3.9
	6	3.4	2.3	4.2	3.2
	7	6.6	8.0	8.5	8.6
	8	.8	1.1	1.5	1.9
	9	3.8	4.0	5.0	5.3
	10	4.2	4.5	6.4	6.1
	11	1.8	2.4	2.7	2.0
	12	3.1	3.6	4.0	3.8
	13	1.4	1.3	2.2	2.6
	14	.8	.9	2.2	3.1
	15	5.5	4.3	7.1	7.0
	16	.1	.7	.7	1.7
	17	2.2	1.4	3.2	3.9
	18	.2	1.3	3.1	4.1
	19	3.2	2.0	5.1	4.1
	20	1.4	2.1	2.8	2.7

From the results given in Table 1, we see that the seven tests in which there is a decided difference in score obtained by the two age-groups (older and younger) are 6, 8, 14, 15, 16, 18, and 19. The younger attain in test 16 (drawings from memory) 343 per cent of the score of the older, in test 18 (dissected sentences) 164 per cent, in test 8 (arrangement of weights) 130 per cent, and in test 14 (three words in one sentence), 133 per cent. In test 6 (repetition of sentences), on the contrary, the corresponding figure is but 72 per cent, in test 15 (comprehension of questions) 90 per cent, and in test 19 (definition of abstract terms) 74 per cent. These percentages must not be taken to mean too much because thus far we have paid no attention to character of disease, but have grouped all psychotic cases together indiscriminately. We have, of course, an entirely different proportion of some mental diseases, such as senile dementia and arteriosclerosis, in the two groups. The effect of this factor, therefore, may be cutting across the effect of chronological age.

The effect of type of disease can be eliminated only by considering each disease separately. Unfortunately we have only one diagnosis where our cases are sufficiently numerous to warrant such comparison. That disease is dementia praecox. We tried many groupings of total score and of chronological age in these cases and arrived finally at seven subgroups each of which had the same range of total score but different chronological ages and each of which contained at least 10 cases under each age-group. The number of cases in some of these age-groups was as great as 52. The subgroups are: Score 46-71, age 10-19 and 20-29; Score 46-71, age 10-29 and 30-49; Score 46-71, age 10-39 and 40-69; Score 72-85, age 10-29 and 30-49; Score 72-100, age 20-29 and 30-39; Score 72-100, age 10-29 and 30-49; and Score 72-100, age 10-39 and 40-69. Score 18-71 gives several groups but is too great a range for reliable comparisons. The comparison of the first twelve tests in the scale reveals nothing decisive, for now a younger, and now an older, group appears to be slightly superior. When, however, we come to the thirteenth test (words in three minutes) we find that in 5 of the 7 sub-groups the younger are uniformly better. The younger ages with entire uniformity give higher scores in tests 14 (three words in one sentence), 16 (drawings from memory), and in 18 (dissected sentences). Their score is greatest in the 16th test (about 200 per cent of the score of the older persons). The older ages, on the other hand, are superior in tests 15 (comprehension of questions), 17 (detection of absurdities) and 19 (definition of abstract terms). Their best performance is in test 17 (about 125 per cent). We have not given the complete data in these cases because it does not seem worth while in view of the relatively small number of cases in the groups and in view of the fact that with dementia praecox we may not be dealing with a unity after all.

The general result is that judging from our psychotic cases alone and given equal general intelligence, patients over 50 years of age excel those under 50 in comprehension of questions, detection of absurdities, and definition of abstract terms; while they are inferior to the younger patients in construction of sentences and drawings from memory.

Comparison of Older Psychotic and Older Normal Persons (Groups a and d)

In the preceding section we have been careful to insist that we have been dealing with results on psychotic patients and that it may not be justifiable to infer from them to normals. The reader may also be tempted to argue that the changes in score which appear with advancing chronological age may be due, not to the influence of age itself, but to the influence of the length of time which a patient has suffered from mental disease. That is to say, it is to be expected that deterioration will continue with the progress of a mental disease. We may have been measuring simply the influence of the course of the disease. If so, the psychotic older persons should differ markedly from normal persons of the same age. The next step, therefore, will be to compare older psychotic persons with their normal contemporaries.

Our groups of psychotic and normal older persons (groups *a* and *d*) are of approximately the same social level, with the former holding the possible advantage. In other regards the two groups are also very similar.

If we consider the groups as a whole and disregard both total score and type of disease, we find that the normal old have a higher average total score than the psychotic old, and that in most cases they have a higher average score on each of the tests. The exceptions are test 13 (words in three minutes) where the averages are the same, and test 14 (three words in one sentence) where the older psychotic patients surpass the older normal persons by one tenth of a point. This result in part answers the question we raised above in regard to the deterioration of the older psychotic persons. The older psychotic persons do show a lower intelligence rating than their normal contemporaries. Since they also have, if anything, a higher social rating, we must conclude that this greater deterioration is due to the progress of mental disease. A question remains, however. Is their greater deterioration a quantitative difference simply, or have the psychotic persons fallen off in certain abilities while retaining others to a normal degree? And further, are our results true for all diseases or merely for their average?

The last question may be answered first and most easily. Our cases include dementia praecox, unclassified paranoid psychosis, syphilitic psychoses, acute alcoholic and deteriorating alcoholic psychoses, manic-depressive insanity and senile dementia. If we consider these diseases separately and compare the results for one with the average attainments of normal older persons, we find that the dementia praecox, senile dementia, and arterio-sclerotic persons differ most from normal persons of their age, while manic-depressive and unclassified paranoid cases differ least. There is no disease, however, which gives results contradictory to our earlier figures for the two groups as wholes.

We come now to the question of the particular tests in which these changes are most marked. To reduce the effect of total score, we must once more fractionate our cases accordingly. We then find two groups with sizes sufficiently great for reliable comparison; those with total scores between 53 and 60, and those with total scores between 76 and 79. From the results thus arranged it appears that the low rating of the psychotic cases in some of the tests, at any rate (such as words in three minutes, three words in one sentence, and drawings from memory) is not due to the fact that the patients are psychotic, for normal persons of the same general age, and of the same mental rating, do *even more poorly in these particular tests than do psychotic persons.*

If, then, the particular failures are due not to the disease but to chronological age, we should get the same general results from the consideration of normal cases.

Comparison of Older Normal and Younger Normal Persons (Groups a, b, and c)

The results for our normal older persons are briefly summarized in Table 2.

TABLE 2

AVERAGE POINT SCALE SCORES OF OLDER NORMAL PERSONS

Chronological Age		50–59	60–69	70–79	80–89
Number of Cases		55	34	13	4
Test	1	3.0	3.0	3.0	3.0
	2	3.8	3.7	3.7	3.0
	3	2.9	2.9	2.9	3.0
	4	3.8	3.3	3.3	3.5
	5	3.8	3.7	3.7	3.3
	6	4.1	3.9	3.4	3.0
	7	7.8	7.3	7.2	8.0
	8	1.8	1.8	1.6	2.0
	9	5.5	5.6	5.2	3.0
	10	5.9	5.6	5.0	4.0
	11	2.6	2.5	2.3	1.8
	12	3.7	3.1	3.4	3.3
	13	1.6	1.9	.8	1.8
	14	1.3	1.0	.7	.5
	15	6.7	5.8	5.5	4.8
	16	.9	.3	.1	.0
	17	3.4	2.9	3.1	2.3
	18	2.6	2.3	.7	.5
	19	4.4	3.5	2.9	1.5
	20	2.3	2.2	1.3	1.0
Total Score		71.9	66.3	59.8	53.3

From this table we see that in general the total score falls off with advancing chronological age among normal persons. Of greater importance than the decrease itself, however, is the *manner* of the falling off, that is to say, the particular tests in which lowering of score first appears or in which it appears in greatest degree. In the table, the small number of cases in the highest age-group makes their average scores unreliable. We shall, therefore, confine ourselves to the three younger age-groups. In these three groups it is evident that with

TABLE 3

AVERAGE POINT-SCALE SCORES FOR NORMAL SUBJECTS OF THREE AGE GROUPS

Total Score	Chron. Age	Test									
		1	2	3	4	5	6	7	8	9	10
46– 52	10–19 (c)	3.0	3.7	2.3	3.4	3.5	?	6.2	.9	4.1	4.6
	20–29 (b)	2.8	3.4	2.9	2.5	3.0	2.4	5.8	1.6	4.7	4.6
	50–84 (a)	2.8	3.6	3.0	3.8	3.2	2.4	6.2	2.0	4.8	5.0
53–60	10–19	3.0	3.7	2.8	3.7	3.8	?	8.2	1.5	4.8	4.8
	20–29	3.0	3.6	2.9	2.9	3.8	2.7	6.5	1.5	5.1	5.0
	50–84	3.0	3.4	2.8	3.3	3.7	3.8	6.6	1.7	4.7	5.1
61–66	10–19	3.0	4.0	3.0	3.9	4.0	?	6.6	1.8	5.1	5.6
	20–29	3.0	3.8	3.0	3.3	3.9	2.9	6.8	1.9	5.4	5.5
	50–84	3.0	3.6	2.8	3.1	3.7	2.8	7.5	1.7	5.2	5.5
67–71	10–19	3.0	4.0	3.0	3.9	4.0	?	6.9	1.6	5.5	5.9
	20–29	3.0	3.7	3.0	3.9	3.9	3.6	6.9	1.9	5.2	5.3
	50–84	3.0	3.7	2.9	3.4	4.0	4.9	7.6	1.7	5.6	5.0
72–75	10–19	3.0	4.0	3.0	3.9	4.0	?	7.2	1.8	5.7	6.2
	20–29	3.0	3.9	3.0	4.2	3.9	4.1	7.1	1.9	5.7	6.2
	50–84	3.0	3.6	3.0	3.8	4.0	5.0	7.5	1.5	5.9	5.6
76–79	10–19	3.0	4.0	3.0	3.9	4.0	?	7.4	1.7	5.8	5.6
	20–29	3.0	3.9	3.0	3.0	3.9	4.5	7.4	2.0	5.6	6.1
	50–84	3.0	3.8	3.0	3.6	4.0	5.6	8.2	2.0	5.4	5.2
80–82	10–19	3.0	4.0	3.0	4.3	4.0	?	7.3	1.9	5.9	6.3
	20–29	3.0	4.0	3.0	4.3	4.0	4.1	7.8	2.0	5.8	6.5
	50–84	3.0	3.8	2.9	3.8	4.0	4.8	8.5	1.5	5.8	6.3
83–100	10–19	3.0	4.0	3.0	4.6	4.0	?	7.9	1.8	5.8	6.9
	20–29	3.0	3.9	3.0	4.7	4.0	4.9	8.0	1.9	5.4	6.8
	50–84	3.0	4.0	3.0	4.3	4.0	4.5	8.1	2.0	6.0	7.1
	AVERAGE FOR THE ABOVE EIGHT GROUPS										
46–100	10–19	3.0	3.9	2.9	4.0	3.9	?	7.2	1.6	5.3	5.8
	20–29	3.0	3.8	3.0	3.6	3.8	3.7	7.0	1.8	5.4	5.8
	50–84	3.0	3.7	2.9	3.6	3.8	4.2	7.5	1.8	5.4	5.6

advancing chronological age there is a tendency for the scores in each test to fall off. This tendency is most marked in tests 16 (drawings from memory), 18 (dissected sentences), 14 (three words in one sentence), and 13 (words in three minutes). It is least marked; on the other hand, in the very easy tests (1, 2, 3, and 5, aesthetic comparison, missing parts, comparison of lines and weights, and counting backwards) and in test 17 (absurdities).

When we had reached this point in our investigation, we

TABLE 3 (continued

AVERAGE POINT-SCALE SCORES FOR NORMAL SUBJECTS OF THREE AGE GROUPS

Test										No. of Cases
11	12	13	14	15	16	17	18	19	20	
1.6	2.3	1.7	1.2	2.6	1.0	.9	.1	.2	.9	45
1.8	3.4	.9	?	4.4	1.0	1.6	?	.9	1.1	40
2.6	2.8	.4	.0	3.4	.0	.0	.0	2.4	.2	5
1.8	2.9	2.0	1.9	3.2	1.2	1.5	1.4	.7	1.2	61
1.9	3.1	2.2	?	4.8	1.2	2.0	?	1.9	1.0	64
2.4	3.4	.7	.0	4.4	.0	1.9	.1	2.9	.8	9
2.2	3.2	2.6	2.7	3.8	1.7	1.7	1.9	1.5	1.6	39
2.5	3.3	1.7	?	5.2	1.5	2.0	?	2.3	1.1	60
2.8	3.6	.8	1.1	5.1	.2	2.3	1.3	3.5	1.2	13
2.3	3.4	2.8	2.8	4.5	2.1	2.1	2.6	2.3	2.3	31
2.4	3.5	1.9	?	5.7	2.3	2.7	?	3.1	1.5	42
2.6	3.6	2.0	.7	6.0	.7	3.6	1.3	2.3	2.0	7
2.6	3.5	2.6	3.6	5.1	1.7	3.6	3.2	3.0	2.4	21
2.6	3.7	2.2	?	6.1	1.6	3.2	?	3.0	1.8	27
2.6	3.9	2.3	1.2	6.6	.3	3.5	2.0	3.0	2.5	8
2.5	3.7	3.3	3.6	5.3	2.5	3.2	3.5	3.3	2.8	31
2.0	3.5	2.4	?	7.1	1.8	3.6	?	4.1	2.1	34
3.0	4.0	1.6	1.5	6.8	.6	4.4	4.0	4.4	1.8	5
2.5	3.6	3.4	3.8	7.1	2.3	3.9	3.8	3.5	3.2	19
2.3	3.9	2.8	?	6.6	2.5	4.0	?	4.1	2.6	26
2.5	3.7	1.9	2.7	7.8	.3	4.3	3.1	4.8	2.8	13
2.6	3.9	3.8	3.8	7.0	3.2	4.0	5.1	4.8	3.9	69
2.7	3.9	3.1	?	6.9	3.2	4.4	?	4.6	3.9	22
2.8	3.7	2.5	2.8	7.9	2.1	4.8	4.9	5.8	4.6	17
AVERAGE FOR THE ABOVE EIGHT GROUPS										
2.3	3.3	2.8	2.9	4.8	2.0	2.6	2.7	2.4	2.3	316
2.3	3.5	2.2	?	5.9	1.9	2.9	?	3.0	1.9	315
2.7	3.6	1.5	1.3	5.8	.5	3.1	2.1	3.6	2.0	77

realized that to make the study complete we needed records of normal persons under 50 years of age of the same general class as our group of older normal persons. It is regrettable that we had not time to continue our work at the General Hospital in which we obtained our older normal cases. We were obliged to use records obtained by other experimenters from other sources. The only records of persons between 30 and 50 years of age which we could easily consult were those of persons so superior intellectually to the cases above considered that they were incomparable. There is, therefore, a gap at ages 30-50 in our Table (3) which summarizes the results of all our normal subjects, including school-children, young men and older persons.

No grade could be given to the school-children in test 6 because the form of this test has been changed since they took the examination. Neither could any grade be given the young men in tests 14 and 18 because so many of this group were illiterate that these tests were omitted. For this group (*d*), therefore, the groups of total scores in the left hand column should read 46-50; 51-56; 57-61; 62-65; 66-68; 69-71; 72-74; and 75-90. In the last section of the table we have given the averages of the eight total-score groups. This, to be sure, is an "average of averages." Even so, however, it is more significant than would be an average of the original scores. In the first place, we could not use the simple average of scores obtained on each test by each of our age-groups because there were so many more high scores among the younger that they would have appeared to too great an advantage. We, therefore, divided our cases in groups by total score attained. The averages for these groups may then be treated as if they had been "corrected" and as if there were an equal number in each total score group, and our "average of averages" is, then, the average score attained on each test by each age-group, supposing the distribution of total scores for the age-groups to be the same.

From Table 3 it appears that tests 1 (aesthetic comparison), 3 (comparison of lines and weights), 5 (counting backwards), 7 (description of pictures), 9 (comparison of terms), and 20 (analogies) show little or no regular change with advancing chronological age. Tests 2 (missing parts), 4 (memory span for digits), 6 (memory span for sentences), and 10 (definition of concrete terms) show a slight tendency toward decrease with advancing age. Tests 8 (comparison of weights), 11 (line suggestion) and 12 (copying square and diamond) show a slight tendency for the score to increase with advancing years. The other tests show decided tendencies. Tests 13 (words in three minutes), 14 (three words in one sentence), 16 (drawings from memory), and 18 (dissected sentences), all show a marked falling-off in score as the chronological age increases. This is most marked in test 16 (drawings from memory). On the other hand, the remaining tests, 15 (comprehension of questions), 17 (absurdities), and 19 (definition of abstract terms), show an increase in score with increased age. From the results we may conclude that the improvement in the ability to comprehend questions comes fairly early, since our young men are so far superior to the school children and are practically the same as the older persons. The improvement in absurdities and definitions of abstract terms, on the other hand, seems fairly regular. The falling-off in giving words in three minutes seems regular, but in drawings from memory, the score does not show any decided decrease until late, that is, the young men differ but little from the school children, while the older persons are decidedly inferior to the young men. Here again we regret that the lack of sufficient data on persons between 30 and 50 prevents us from more than guessing at the points at which the various abilities show changes.

The main conclusion to be drawn from our work thus far is that whether we study psychotic or normal persons with approximately the same total score, the younger persons tend to excel the older in giving words in three minutes, in building sentences and in drawing from memory; while the older excel the younger in comprehending questions, in detecting absurdities, and in defining abstract terms. Our results were so consistent throughout that we next turned to the literature to see if experimenters who were looking for something else had results which agreed at all with ours.

Results from Previously Published Work

In our paper on "Significant Responses in Certain Memory Tests," (referred to in footnote 2), in which we consider only Psychopathic Hospital cases, some of whom were diagnosed as "not insane," we find no uniform changes with increasing chronological age in the score of memory span for digits. We do find, however, a decided tendency for the score of drawings from memory thus to decrease. The point at which this decrease begins differed for the different mental diseases, but in all it begins before age 50. A similar decrease is found in the scores of memory for short paragraphs.

Our results in general are also confirmed by a study of the contributions of two other writers. Both of these are reports of Binet examinations (one, the Goddard, 1911, and the other, the Stanford revision). The differences in grading and in the tests themselves make a rough comparison the only one possible.

An article by Wender[5] gives much valuable material, though his conclusion on the basis of his selected cases that his data furnish proof for the necessity of a revision of the scale appears unjustified. His Table II shows the tests passed by 30 cases of senile dementia or arterio-sclerosis, with an average chronological age of 74 and an average mental age of 9.4. From it we have calculated the percentage of the cases which passed each test which corresponds to a test on the Point Scale. The results of such calculation are given in Table 4.

TABLE 4

PERCENTAGE OF WENDER'S CASES PASSING TESTS SIMILAR TO POINT SCALE TESTS

Binet 1911 Test	Corresponding Point Scale Test	Percentage Passing
VI, 5	1	87
VII, 3	2	37
VIII, 5; X, 3; XII, 1	4	65
VIII, 2	5	67
VII, 2; XV, 1	7	67
IX, 5	8	13
VIII, 1	9	83
VI, 2; IX, 2	10	70
XII, 4	11	17
VII, 4	12	67
XI, 3	13	23
X, 5; XI, 2	14	27
X, 4	15	57
X, 2	16	7
XI, 1	17	33
XI, 5	18	27
XII, 2	19	83

The most striking points about Table 4 are that test 19 (a 12-year Binet test) should have been passed by 83 per cent of a group whose average mental age was only 9.4 and that test 16 (a 10-year test) should have been passed by only 7 per cent of the same group. Tests 2 and 8 (missing parts and arrangement of weights) also give a low average when we consider the age level to which they presumably belong. Test 8 was one of the tests in which our younger psychotic patients were superior to the older psychotic patients. Test 2 was found to change very little until we reached the 80-89 group of our normal older persons. But, other than this, our results do not seem to uphold Wender's findings for tests 2 and 8. The successes in Wender's group are more striking than the failures. We have already mentioned the remarkable performance in test 19 (definition of abstract terms). Tests 15 and 17 (comprehension of questions and absurdities) also give very high percentages. Test 15, a 10-year test, is passed by 57 per cent of the cases; and test 17, an 11-year test, by 33 per cent. These three tests, it will be remembered, were the three which we found to be particularly easy for the older persons, and we may say, therefore, that Wender's work supports our results in so far as the two have anything in common.

---

[5] "The Applicability of Binet-Simon Tests in Psychoses of the Senium." *N. Y. Medical Journal*, March 6, 1915.

The second contribution mentioned is in Terman's statistical account of the bases of the Stanford revision.[9] From this account we have taken the results on Knollin's unemployed men (hoboes), chronological ages 21-60, but chiefly 25-40, mental ages 10-18. The results are summarized in Table 5.

TABLE 5

PERCENTAGE OF KNOLLIN'S CASES PASSING TESTS AT DIFFERENT MENTAL AGES

Stanford Revision Test	Percentage Passing at Mental Ages								Corresponding Point Scale Test
	10	11	12	13	14	15	16	18	
IX, 2		70	79						8
X, 6		40	60	56	60				13
IX, 5	30	67	75	96					14
X, 5	70	85	98	100	100				15
X, 3	41	52	67	70	73				16
X, 2	50	60	82	81					17
XII, 4		0	27	54	63				18
XII, 2		50	90	83	97				19

The blank spaces in Table 5 seem to mean that if the table were completed, there would be zeros in the blanks to the left of the figures given and hundreds to the right. It is, of course, impossible to determine what these subjects might have done on the other comparable tests of the Point Scale because the arrangement of the Stanford is such that they were not given.[7] Test 18 (dissected sentences) is very difficult for the Knollin group. Although it is a 12-year test, only 27 per cent pass at age, and only 63 per cent of cases with a mental age of 14 pass. Likewise test 16 (drawings from memory) is difficult, for although a 10-year test, only 41 per cent of those with mental age 10 pass it, and only 73 per cent of mental age 14 pass. Test 14, again (three words in one sentence), is done but poorly. On the other hand, test 15 (comprehension of questions), gives percentages which are normal for its place in the scale and test 17 (absurdities) is not far behind. Test 19 (definition of abstract terms) is evidently done better than we should expect from school children. We have, then, roughly, though not as decidedly as for Wender's older cases, the conclusion that tests 14, 16, 18, and probably 13 (three words in one sentence, drawings from memory, dissected sentences, and words in three minutes) are difficult for hoboes of chronological ages between 25 and 40, and that test 19 (definition of abstract terms) is easy for them. Furthermore, if we compare these cases with those for Williams' juvenile delinquents,[8] ages mostly 14 to 21, we find that the hoboes do much better than the delinquents in test 19 (definition of abstract terms), and that the delinquents far excel the hoboes in tests 13, 16, and 18 (words in three minutes, drawings from memory, and dissected sentences). The same tendency holds true if we compare them with still younger cases. The only comparison possible between the business men and the High School pupils[9] is on tests 18 and 19 (dissected sentences and definition of abstract terms). Test 19 is passed by 1 per cent more pupils than business men, but in test 18 this percentage rises to 10, showing that test 18, at least, is harder for business men. That is to say, the falling off in ability to put together dissected sentences has begun already in middle aged business men. In the comparison of abstract terms (a Stanford test which is similar to the Point Scale definition of abstract terms) 13 per cent more business men pass than do pupils. We find again then that younger persons consistently excel older in certain tests and are inferior to them in others.

---

[9] Terman and others, "*The Stanford Revision and Extension of the Binet Simon Scale for Measuring Intelligence*," 1917, p. 163 ff.

---

[7] The above seems to the writers a strong argument for the use of scales of the type of the Point Scale in all cases where there is a probability of wide "scatter." In year scales there is far greater chance of passing over some defect or peculiarity simply because it is not expected at the chronological or mental age of the subject.

[8] Terman and others, *op. cit.*, p. 170 ff.

[9] *op. cit.* p. 171 ff.

The results of other investigators, therefore, although obtained for quite different purposes, agree almost absolutely with our findings as to the influence of chronological age.

Probable Reasons for the Differences Between Older and Younger Persons

The above data show with reasonable certainty that there are decided changes in the distribution of abilities (as shown by the Point Scale) as persons get older. These changes appear whether we compare the insane or the normal, whether we lump all our cases together or fractionate them by particular disease, whether we compare the very old with the very young, or whether we compare the middle-aged with young adults. The only condition which must be observed is this one: that only those whose total score, or general level of intelligence, is approximately the same may be reliably compared. If we do not make this restriction, the young will be found to excel in each test as well as in the total score.

The next question is: Why do we find this change in ability as persons get older? It is evident, of course, that as a person grows older he loses some of the abilities he had as a child or young adult. We are all familiar with aged persons who fail to remember what happened yesterday, but who expect their grandchildren to recall, as they themselves do, events which happened many years ago. We say commonly that old people have poor recent, but good remote memory, but we seldom inquire into their abilities outside the field of memory.

In the present paper we have tried to determine more exactly the abilities and disabilities of the old and to estimate roughly, at least, the age at which changes are most marked. It is to be supposed that a defect in memory which is clearly present in a person of 80 must have been coming on for some time.

We have found that in the abilities which are tested in the Point Scale the old persons have deteriorated much more in some than in others. The reasons for the individual losses we conceive to be three. In the first place, there is some actual loss in ability. This is shown particularly in the drawings from memory. From our practical acquaintance with aged persons we are forced to conclude that they are actually unable to recall certain recent impressions. In the second place, there is a lack of practice in certain kinds of performance. Such, for example, we take to be the case in the construction of sentences. There seems to be a possibility that, given sufficient incentive and sufficient practice in this test, an older person may equal the performance of his juniors. The difficulty is that the common incentives such as praise, approval, etc., which are so effective with children, are of little avail with the old. We come now upon the third point which is probably the key of the whole problem. The younger subjects are almost invariably more alert and interested. Their experience is such that they fit more naturally into the test situation. They appear more adaptable than the older ones. Moreover, the tests in which they excel are those which most resemble " stunts " or " puzzles " and which, therefore, require not only willingness, but also a rapid adjustment of the subject. If we consider the tests in which the older subjects are superior we find them to be the ones which are more like the problems which arise in the daily life of adults and which could be answered best by persons who had had the accumulated experience of years.

There seems, therefore, to be such a decided break between the older and the younger persons that it is not fair by the former to grade them by an examination intended primarily for adolescents. The whole question of the applicability or fairness of any such examination to older subjects therefore depends on the purpose for which it is given.

The purposes of examinations of persons over 50 seem to us to be two: first, the determination of the degree of deterioration or aberration present; second, the determination of the presence of feeble-mindedness. In both cases there are two possible standards of comparison, namely their supposed former ability (or that of average normal young person) and the average ability of their contemporaries. It is without doubt interesting to note that a person who once had a mental age of 18 has now one of only 10. It is, however, of much greater importance to know whether the average person of the same present chronological age and of the same former mental age has deteriorated to the same degree. If the patient's deterioration is the same in amount and kind as that of his normal contemporary, then we cannot lay that deterioration to the presence of mental disease or to initial feeble-mindedness. Moreover, if the history of an old person convinces us that he has always been of a low grade mentally it is often desirable to know the maximum mental age which the patient ever attained. Our best guess here would be based on those tests in which the normal contemporary has not shown deterioration.

In most work with psychological examinations we have the constant difficulty that many non-psychological persons (and, alas! some psychologists) take our results as too simple. They read the mental age, without taking any notice of the comments of the examiner. If the mental age is less than 12 they glibly diagnose the patient feeble-minded. In order to circumvent such hasty diagnosticians and in order to give a mental age which shall more exactly express the ability of the subject before the influences of old age became marked, we have calculated some allowances which should be made in the case of persons over 50 years of age.

Suggested Score Corrections for Old Age

Already at this hospital we had been in the habit of making allowance for the omission of certain tests and we now applied the same method to a scheme for discounting the effect of advancing years. Perhaps it will be as well to give the history of the previous work, so that the present calculations will not seem too fanciful. To be sure, the plan we are about to present has obvious faults and we can claim for it no more than fairly satisfactory results. We give it here in the hope that the idea will lead some others to similar work and will in the end result in an accurate and theoretically correct table.

The first problem of the kind which arose was the question of how to grade patients who were totally deaf and who, therefore, could not be given tests 4 and 6 (auditory memory span for digits and sentences). Our procedure at first was to add the scores with these tests omitted, call that the minimum mental age, then add to that the highest score obtainable on the two omitted tests, call that the maximum mental age, and then say that the true mental age lay somewhere between those two limits. This was fairly satisfactory, but we thought it possible to get a more accurate statement. This we computed from our table of scores for each test which were to be expected for different ranges of total score.[10] From the table we calculated the amount of credit to be expected on tests 4 and 6 for each of the ranges of total score. We then constructed a table giving the amount that should be added for each total score obtained when the two tests were omitted.

---

[10] This table was published with some printer's errors, (later corrected) in the *Journal of Abnormal Psychology*, XIII, 1918, p. 77.

We later made similar tables of corrections for omission of tests 14 and 18 (lack of education) and for tests 1, 2, 3(a), 7, 11, 12, 16, and 18 (total blindness). The corrections for lack of education were adopted by the Division of Psychology in the Surgeon General's Office for use in the examination of illiterates. The corrections are given in Table 6.

TABLE 6

CORRECTIONS FOR POINT SCALE NORMS WHEN CERTAIN TESTS ARE OMITTED

When 4 and 6 are Omitted (Deafness)		When 14 and 18 are Omitted (Education)		When 1, 2, 3 (a), 7, 11, 12, 16, and 18 are Omitted (Total Blindness)	
For Scores:	Add:	For Scores:	Add:	For Scores:	Add:
13–25	5	18–51	0	7–13	11
26–60	6	52–58	2	14–15	15
61–78	7	59–62	4	16–21	16
79	8	63–69	6	22–28	17
80–91	9	70–74	8	29–34	18
		75–77	9	35–39	21
		78–90	10	40–42	24
				43–48	27
				49–50	29
				51–52	30
				53	32
				54	33
				55–66	34

With these tables as models, we proceeded to make a similar table to correct for advanced chronological age. We have found throughout, as we have said, that the older subjects are almost without exception poorer in tests 13, 14, 16, and 18 than younger persons attaining the same total score. We have therefore supposed that these tests should be omitted in giving the examination to old people and have calculated the corrections for such omission. We do not mean that the tests should actually be omitted. On the contrary, if a person of over 50 years of age obtains a high score on the four tests, it is evident that he has not begun to lose certain abilities which many of his contemporaries have lost. In other words, in our opinion failure on tests 13, 14, 16, and 18 on the Point Scale means little or nothing if the subject is advanced in years, while success on those tests may be very significant. The corrections which we offer tentatively for this group of advanced ages are given in Table 7.

TABLE 7

CORRECTIONS FOR POINT SCALE NORMS TO BE USED WITH OLDER SUBJECTS

When Tests 13, 14, 16 and 18 are Omitted
(Advanced Chronological Age)

For Scores:	Add:
18–36	0
37–43	1
44–48	3
49–53	5
54–55	9
56–58	12
59–61	13
62–66	15
67–69	17
70–82	18

At first thought it may appear that if we correct for failures which seem to be due to advanced age alone, we should also correct for successes which are apparently due to the same cause. Perhaps we should. If the idea were carried to its logical extreme we would be correcting for every test except 1 and 20, the only ones in which the average score for young and old is identical. Such a procedure would, of course, be meaningless, and would amount to giving a mental age on the basis of two tests alone. Somewhere, then, we must draw the line between no correction and total correction. We considered at first correcting for those tests in which one age gave an average score which was 120 per cent of the score obtained by the other age. This limit, however, would make us correct for 8 tests, in five of which the younger and in three of which the older were superior. Eight seemed such a large percentage of the total number (20) of tests that we were afraid we were again basing mental age on too few tests. If the limit were raised to 200 per cent, we would be correcting for only three tests, 13, 14, and 16. Test 18 which came next on the list with the younger excelling the older by 136 per cent, was later included because the test is one which many of the older subjects dislike, and which they often cannot see to read. The actual limit used was, therefore, 136 per cent.

Conclusions

1. There are certain definite changes in the distribution of scores on the Point Scale as the chronological age of the subject increases.

2. These changes are evident in both normal and psychotic persons.

3. There are three probable reasons for the changes: loss of ability, lack of practice, and absence of alertness or of interest in the older subjects.

4. The mental condition of a subject over 50 years of age will be much more accurately presented if two mental ages are given: one which compares him with his own adolescent ability (or with that of normal young persons), and one which compares him with his normal contemporaries.

5. A mental age which compares a subject with his normal contemporaries may be calculated from our Table 7.

SOME RESULTS AND INFERENCES DERIVED FROM THE USE OF THE ARMY TESTS AT THE UNIVERSITY OF MINNESOTA[1]

By M. J. VAN WAGENEN, University of Minnesota

During the two academic years 1917-18 and 1918-19 the Army Test Alpha or its equivalent, Form E, was given to the freshman classes of nearly all the colleges at the University of Minnesota. In some of the colleges the results were used as an aid in diagnosing the causes of student failures. In other colleges the purpose was purely an experimental one. The use of the tests has revealed some significant information regarding the students of the various colleges. For instance, over eighty per cent of the student body of the University as a *whole were found to come from the upper fifteen to twenty* per cent of the population in general. With the exception of two of the individual tests the women of the several colleges did just as well or even better than the men of the same colleges, but in these two tests—the range of information test and the arithmetic problems test—from seventy to seventy-five per cent of the men did as well or better than the median woman. Making correction for the excess of overlapping due to the use of a single test, from sixty-five to seventy per cent of the men may be expected to do as well or better in solving arithmetic problems and to have as wide or a wider range of the kind of information called for in the Army Tests than has the median woman of the same college. *Incidentally, the students of the College of Education* were found to be a markedly superior group,—corresponding well with the highly selected group in the Medical School,—a fact which speaks well for the leadership of the teaching profession. More surprising was the discovery that eighty per cent of the students of one technical college fell below the median student of another technical college in the abilities involved in the doing of such tasks as those making up the Army Test. And this is evidently a stable condition, for the differences of the second year were practically the same as they were for the first year the test was given.

---

[1] Read before Section L of the American Association for the Advancement of Science at the Annual Meeting, held in St. Louis, December, 1919.

Interesting and valuable as such information may be to those in charge of the various colleges, nevertheless we are more concerned at the present time with the actual value of the Army Test and similar mental tests as a basis for predicting what kind of academic work will be done by each student during his college course, and especially with the further use that may be made of such tests as a basis of acquiring more definite knowledge of vocational aptitudes than merely the settling of the question as to whether or not a student is likely to succeed in completing a college course.

During 1918-19 one of the graduate students in the department of educational psychology, Miss Judith Jacobs, worked out the correlations of the scores of the Army Test and also the correlations of the scores of each of the individual tests with the averages of the marks made by the women students of the College of Science, Literature and the Arts during each semester of their freshman year. The second Army Test given to the University freshmen was also given to a large number of sophomore women, most of whom had taken a similar test, Form E, the previous year. This made possible a comparison of the students' scores on the two tests taken some fourteen months apart and also the determination of the relation of each of the two Army Tests used to the averages of the academic marks made by the women during each semester of their freshman year. It was necessary to confine this study to the records of the women students of the College of Science, Literature and the Arts, as no other college offered so large a group of records for the women and the work of the men had been too badly broken up with the advent of the S. A. T. C. to permit of any accurate study being carried on with the use of their academic marks.

For 139 women students between the ages of eighteen and twenty, Miss Jacobs found a correlation of +.39±.05 between the scores made in the Army Test, Form E, and the averages of the marks for the first semester of the freshman year, a correlation of +.35 ±05 with the averages of the marks for the second semester, and a correlation of +.34±.05 between the scores made in the Army Test Alpha, Form 6, and the averages of the marks for the first semester of the previous (freshman) year. Miss Jacobs also found that the correlation between the averages of the marks for the first semester of the freshman year and those of the second semester was +.78±.02, while the correlation between the Army Test, Form

E, and Alpha, Form 6, was +.82 ±.02. The correlation between the scores of the Army Test, Form E, and the averages of the marks for the first quarter of the sophomore year dropped to +.13 ±.05. At the same time the correlations between the averages of the marks for the first and second semesters of the freshman year and the averages of the marks for the first quarter of the sophomore year fell to +.51 ±.05 and +.52 ±.05 respectively. The very low correlation between the Army Test scores and the averages of the students' marks for the first quarter of the sophomore year is undoubtedly due to a very large extent to the errors of judgment in the assignment of the sophomore marks, since the quarter involved less than ten weeks of academic work. If this assumption is true then any increase in the accuracy of the judgments of the instructors regarding the quality of students' achievements should result in higher correlations between the Army Test scores and the averages of the academic marks. Later it will be shown that this was exactly the condition found.

For the individual tests of the Army Test the averages of six correlations of the scores with the averages of the marks of one semester are as follows:

Practical judgment scores	and averages of marks for 1 semester	r = —.09
Oral Direction scores	" " " " " " "	r = —.01
Series completion scores	" " " " " " "	r = +.15
Disarranged sentences scores	" " " " " " "	r = +.21
Analogies scores	" " " " " " "	r = +.25
Synonym-antonym scores	" " " " " " "	r = +.27
Range of information scores	" " " " " " "	r = +.31
Arithmetic problems' scores	" " " " " " "	r = +.36

The scores of two of the eight individual tests show practically no relationship to averages of the marks, while the scores of five show a correlation above +.20, the scores of two of them —the range of information test and the arithmetic problems test—showing a correlation above +.30.

The relationships between the scores of the two Army Tests, those between the averages of the marks for the two semesters of the freshman year, and those between the scores of the Army Tests and the averages of the marks for each semester are shown in the following table more clearly perhaps than they are expressed by the coefficients of correlation. By first finding the standard deviation or sigma for the scores of the Army Tests and for the averages of the marks, and then finding in which of the ten one-half sigma units each individual score fell, we have the table below. It reads: Of the 141 students who took both Army Tests, 39 per cent stood in the same one-half sigma unit or differed in their standing in the two tests by less than one-tenth of the range of the scores of the college students; 41.9 per cent differed by not more than one one-half sigma unit or by less than one-fifth of the range of the scores; 15.6 per cent differed by not more than two one-half sigma units or by less than three-tenths of the range of the scores; 2.8 per cent differed by not more than three one-half sigma units or by less than four-tenths of the range of the scores; while only 0.7 per cent differed by as much as four one-half sigma units.

TABLE I

Amount of Difference in one-half Sigma Units	Tests form E and Form 6	Averages of Grades of 1st Semester and of 2nd Semester	Form E and Averages of Grades of 1st Semester	Form E and Averages of Grades of 2nd Semester	Form 6 and Averages of Grades of 1st Semester	Form 6 and Averages of Grades of 2nd Semester
No. of Cases	141	224	241	235	105	105
0 Difference	39.0%	34.4%	19.5%	19.2%	15.2%	14.3%
1 "	41.9%	37.5%	30.7%	28.5%	39.1%	38.1%
2 "	15.6%	18.7%	17.4%	19.6%	21.9%	22.9%
3 "	2.8%	7.1%	16.6%	15.7%	13.4%	17.1%
4 "	0.7%	0.4%	11.2%	11.5%	8.6%	4.8%
5 "		0.9%	4.2%	3.8%	0.9%	2.8%
6 "		0.9%	0.4%	1.7%	0.9%	

Although two of the individual tests of the Army Test contribute practically nothing and a third but little toward prognosticating the academic record of a student, it is nevertheless astonishing how much the Army Test, taking less than forty-five minutes to give, will tell about a student's probable success in making academic marks during either semester of the freshman year. Over fifty per cent of the students will not change their standing in academic marks from their standing in the Army Test by more than one one-half sigma unit or by more than one-fifth of the range of the scores of college students or by more than a change from a D to a C or from a C to a B, as these marks are given to college students. Over twenty-five per cent of the students will make a greater change in their standing in their academic marks from the first semester to the second, even when the same subjects are studied and, for the most part, with the same instructors. As far as making academic marks in either semester of the freshman year is concerned less than fifteen per cent of the students will be misplaced by the Army Test by more than three one-half sigma units or by four-tenths of the range of the marks or by as much as a change from an E to a C, or from a D to a B, or from a C to an A. True it is that this is a rather large misplacement, but over two per cent of the students are displaced just as much in their academic marks from the first semester of their freshman year to the second semester. Then, too, were the academic marks of both semesters of the freshman year taken into account the amount of misplacement by the Army Test would be considerably less than here indicated.

Equally significant with the determination of the amount of change that occurs in connection with the Army Tests and college marks is the determination of the direction of the changes that may be expected of students of different degrees of ability. For this study three groups of students were selected: those who stood in the second and third one-half sigma units or in the E and D— units, those who stood in the fifth and sixth one-half sigma units or in the C— and C units, and those who stood in the eighth and ninth one-half sigma units or in the B and A— units. The units referred to are the one-half sigma units or the respective tenths of the base line of a normal surface of frequency lying between —2.5 sigma and +2.5 sigma.

The first table below (Table II) and subsequent tables (Tables III to VII) read: Of the 23 students who stood in the second and third one-half sigma units from the bottom of the distribution or in the E and D— group in the first Army Test, 34.8 per cent did not change their position in the second Army Test as much as one one-half sigma unit or as much as the difference between an E and a D—; 39.1 per cent changed as much as one one-half sigma unit but not as much as two one-half sigma units or not as much as the difference between an E and a D; 21.7 per cent changed as much as two one-half sigma units but not as much as three one-half sigma units; while only 4.4 per cent changed as much as four one-half sigma units or as much as from an E to a C. While 34.8 per cent of the group did not shift as much as one one-half sigma unit in either direction, 52.2 per cent made a higher position in the second test than in the first one and 13 per cent made a lower position in the second test. Of the group of fifty-seven C— and C students 36.8 per cent remained practically stationary in the second Army Test, while 24.5 per cent made a higher position and 38.6 per cent made a lower position. Of the 15 B and A— students 53.3 per cent remained in the same positions, 26.7 per cent attained a higher position and 20 per cent dropped to a lower position.

TABLE II

AMOUNTS OF CHANGE OF POSITION FROM ARMY TEST, FORM E, TO ARMY TEST ALPHA, FORM 6, IN THE CASE OF THREE GROUPS OF STUDENTS

Amount of Difference in One-half Sigma Units	2nd and 3rd one-half Sigma Units or the E and D—Group of Students	5th and 6th one-half Sigma Units or the C—and C Group of Students	8th and 9th one-half Sigma Units or the B and A—Group of Students
No. of Cases	23	57	15
0 Difference	34.8%	36.8%	53.3%
1 "	39.1%	45.6%	33.5%
2 "	21.7%	15.8%	6.6%
3 "		1.7%	6.6%
4 "	4.4%		
Per Cent of Cases Falling in Higher Units in 2nd Test	52.2%	24.5%	26.7%
Per Cent of Cases Falling in Lower Units in 2nd Test	13.0%	38.6%	20.0%

The table (Table III) giving similar facts for the averages of the academic marks of the two semesters of the freshman year shows the same tendencies; namely, the smallest amount of shifting occurs among the superior students and it is just as likely to be a shift to a higher as to a lower position. The greatest amount of shifting occurs among the students making the lower records and is more likely to be a shift upward than a shift downward. The dominant shifting tendency among the students of the middle group is a downward one in the second semester's work as well as in the second Army Test.

TABLE III

AMOUNTS OF CHANGE OF POSITION FROM MARKS OF THE FIRST SEMESTER TO THE MARKS OF THE SECOND SEMESTER IN THE CASE OF THREE GROUPS OF STUDENTS

Amount of Difference in one-half Sigma Units	2nd and 3rd one-half Sigma Units or the E and D—Group of Students	5th and 6th one-half Sigma Units or the C—and C Group of Students	8th and 9th one-half Sigma Units or the B and A—Group of Students
No. of Cases	39	84	22
0 Difference	15.4%	39.3%	36.4%
1 "	46.2%	39.3%	45.4%
2 "	23.1%	17.8%	13.6%
3 "	12.8%	2.4%	
4 "			
5 "	2.5%	1.2%	
6 "			4.5%
Per Cent of Cases Falling in Higher Units in 2nd Semester	53.8%	22.6%	31.8%
Per Cent of Cases Falling in Lower Units in 2nd Semester	30.8%	38.1%	31.8%

The next four tables (Tables IV to VII) indicate in the same way the degree of stability and the amount and direction of the shifting of positions from those achieved in the Army Tests to those made in the academic grades.

TABLE IV

AMOUNTS OF CHANGE OF POSITION FROM THE ARMY TEST, FORM E, TO THE MARKS OF THE FIRST SEMESTER OF THE FRESHMAN YEAR

Amount of Difference in one-half Sigma Units	2nd and 3rd one-half Sigma Units or the E and D— *Group of Students*	5th and 6th one-half Sigma Units or the C— and C *Group of Students*	8th and 9th one-half Sigma Units or the B and A— Group of Students
No of Cases	48	84	30
0 Difference	29.2%	19.0%	13.3%
1 "	22.9%	39.3%	23.3%
2 "	8.3%	13.1%	20.0%
3 "	14.6%	14.3%	23.3%
4 "	16.7%	13.1%	20.0%
5 "	8.3%	1.2%	
Per Cent of Cases Falling in Higher Units in First Semester's Marks	60.4%	35.7%	6.7%
Per Cent of Cases Falling in Lower Units in 1st Semester's Marks	10.4%	45.2%	80.0%

TABLE V

AMOUNTS OF CHANGE OF POSITION FROM THE ARMY TEST, FORM E, TO THE MARKS OF THE SECOND SEMESTER OF THE FRESHMAN YEAR

Amount of Difference in *one-half Sigma Units*	2nd and 3d one-*half Sigma Units* or the E and D— Group of Students	5th and 6th one *half Sigma Units* or the C— and C Group of Students	8th and 9th one *half Sigma Units* or the B and A— Group of Students
No. of Cases	47	82	28
0 Difference	12.8%	20.7%	7.1%
1 "	29.8%	34.1%	28.6%
2 "	19.1%	23.2%	21.4%
3 "	10.6%	13.4%	17.8%
4 "	21.3%	7.3%	10.7%
5 "	4.3%	1.2%	14.3%
6 "	2.1%		
Per Cent of Cases Falling in Higher Units in 2nd Semester's Marks	68.1%	34.1%	10.7%
Per Cent of Cases Falling in Lower Units in 2nd Semester's Marks	19.1%	45.1%	82.1%

TABLE VI

AMOUNTS OF CHANGE OF POSITION FROM THE ARMY TEST ALPHA, FORM 6, TO THE MARKS OF THE FIRST SEMESTER OF THE FRESHMAN YEAR

Amount of Difference in one-half Sigma Units	2nd and 3rd one-half Sigma Units or the E and D— Group of Students	5th and 6th one-half Sigma Units or the C—and C Group of Students	8th and 9th one-half Sigma Units or the B and A— Group of Students
No. of Cases	14	35	20
0 Difference	21.4%	14.3%	20.0%
1 "	21.4%	45.7%	35.0%
2 "	15.4%	25.7%	10.0%
3 "	21.4%	8.6%	15.0%
4 "	21.4%	5.7%	15.0%
5 "			
6 "			5.0%
Per Cent of Cases Falling in Higher Units in 1st Semester's Marks	71.4%	45.7%	0.0%
Per Cent of Cases Falling in Lower Units in 1st Semester's Marks	7.1%	40.0%	80.0%

TABLE VII

Amounts of Change of Position From the Army Tests Alpha Form 6, to the Marks of the Second Semester of the Freshman Year

Amount of Difference in one-half Sigma Units	2nd and 3rd one-half Sigma Units or the E and D— Group of Students	5th and 6th one-half Sigma Units or the C— and C Group of Students	8th and 9th one-half Sigma Units or the B and A— Group of Students
No. of Cases	14	36	19
0 Difference	7.1%	27.8%	0.0%
1 "	42.9%	36.1%	36.8%
2 "	15.4%	22.2%	26.3%
3 "	21.4%	8.3%	21.1%
4 "	7.1%	5.6%	5.3%
5 "	7.1%		10.5%
Per Cent of Cases Falling in Higher Units in 2nd Semester's Marks	78.6%	36.1%	21.1%
Per Cent of Cases Falling in Lower Units in 2nd Semester's Marks	15.4%	36.1%	78.9%

The facts of these tables may be briefly and conveniently summarized as follows: the per cents of students making low scores in the Army Tests who stay in the same positions in their academic marks are relatively low, ranging from 7 per cent to 29 per cent in the four tables. The per cents of low standing students making lower positions in academic marks than in the Army Tests are also small, ranging from 7 per cent to 19 per cent; while the per cents making higher positions in academic marks than in the Army Tests are relatively high, ranging from 60 per cent to 78 per cent. In the case of the middle group of students—the C— and C students—from 35 per cent to 45 per cent are likely to achieve a higher position in academic grades than in the Army Tests. Likewise from 35 per cent to 45 per cent are likely to fall to a lower position in the academic grades than in the Army Tests. The amount of shifting in this group is smaller, however, than in either the superior or the low standing group of students. Not more than from 15 per cent to 30 per cent of the average students shift their positions in academic marks from those made in the Army Tests by more than the difference between a D and a C or that between a C and a B. In the case of the superior students just the opposite tendency is found from that found among the students making the lower records in the Army Tests. Practically 80 per cent of those making the higher positions in the Army Tests make somewhat lower positions in their academic grades, while only from 0 to 20 per cent achieve higher positions in their academic marks than in the Army Tests.

This situation is highly significant in dealing with the use of mental tests for prognosticating a student's probable academic success. It would seem that the tests prove least useful just where reliable results from their use is most needed; namely, in eliminating those most likely to fail in their college work and in selecting for special groups those who are most likely to attain the higher degrees of success. When it is recalled that of those standing low in the Army Test *far more of them are likely to reach only slightly higher* positions in their academic work it is not improbable that *the extra amount of effort put into their work by the very* low standing students may account in part for the large percentage tending to attain higher positions in their academic work. Likewise the fact that the students standing in the higher positions in the tests tend to fall in somewhat lower positions in their academic work may in part be accounted for by a lack of maximum application to their work, especially when there is not the external pressure to keep them working at their maximum effort that is exerted upon the less capable students. In addition to these factors some amount of shifting is undoubtedly to be accounted for by the large amount of outside work done by not a few of the students, and also by the insistent demand of the more serious social activities upon the time and effort of the more socially inclined individuals. It is not improbable, too, that it is particularly difficult for the instructors to select the exceptionally gifted college students and to accurately gauge their abilities and their achievements, just as for the elementary school teacher to select the exceptionally gifted pupils of her class has been found to be attended with great inaccuracy.

There is yet, however, far too large a discrepancy between the positions attained in the Army Tests and those achieved in academic marks to warrant the use of the Army Tests for purposes of rigid selection. Two questions insistently arise at this point: First, are mental tests, such as the Army Tests, improvable to the degree that their use will be feasible in selecting and classifying college students? To this question an affirmative answer is undoubtedly warranted at the present time. Second, is it possible to improve the tests to the point of reducing the discrepancies within reasonable limits without at the same time securing more adequate measurements of the actual achievements of college students than is now afforded by the academic marks, which are based very largely upon the subjective judgments of the instructors? An affirmative answer to this question is reasonably doubtful, and this, too.

is one of the more important conditions tending to minimize the actual value of the use of mental tests in connection with the admission and classification of college students.

Mental tests such as the Army Test are improvable for the end in view in two directions: first, additional tests showing a fair degree of correlation with academic marks and at the same time contributing elements not found in the tests already available may be included with those now showing some degree of relationship to academic marks; second, the reliability of the tests may be increased by improving the tests and by making them more extensive so that similar series of tests will give more nearly the same arrangements of scores for the individuals tested. The fact that nineteen per cent of the subjects tested by the two Army Tests showed a change of one-fifth or more of the total range of the scores for college students is very clear evidence of the need of more extensive tests—tests lasting a few hours instead of a few minutes. This is not a criticism of the Army Test for the purpose for which it was intended. To classify the upper twenty per cent of the population in small units is a much more difficult task than to classify the total population into relatively large units. Incidentally, one of the strangest phenomena in connection with the use of mental tests has been the demand of so many intelligent people in the field of education that the tests accomplish in a very few minutes what they themselves willingly admit cannot be as well accomplished by any other means in any length of time.

Just as increasing the length of the tests, other conditions being equal, increases the reliability of the tests, so increasing the number of judgments regarding the academic work of the students ought, other conditions being equal, to increase the reliability of the marks, and, at the same time, enhance the relationship between the tests and the academic marks if errors of judgment regarding academic achievements are in part responsible for the relatively low correlations between the test scores and the averages of the academic marks. In the case of eighty-four sophomore women for whom complete records were available the following coefficients of correlation were obtained when the scores made in the two Army Tests, Form E and Alpha Form 6, were correlated with the averages of the academic marks for both the first and the second semesters.

Army Test, Form E and averages of marks for the two semesters $r = +.46 \pm .06$.

Army Test Alpha, Form 6 and averages of marks for the two semesters r=+.50.

For the five individual tests of the Army Tests that gave the highest correlations with the averages of the marks, the following average correlations were obtained with the averages of the marks for both semesters:

Disarranged sentences scores	and av. of marks for both semesters	r = +.33
Synonym-Antonym scores	" " " " " " "	r = +.35
Range of information scores	" " " " " " "	r = +.36
Analogies scores	" " " " " " "	r = +.37
Arithmetic problems 'scores	" " " " " " "	r = +.41

When the scores for these five tests were combined and the correlations between the two combinations and the averages of the marks for the two semesters obtained the following coefficients were found:

Five tests of Form E and averages of marks for both semesters r = +.44
Five tests of Alpha Form 6 and av. of marks for both semesters r = +.47

From these results two facts are evident: the correlations between the Army Test scores and the academic marks were very noticeably raised—approximately one-tenth—when the number of academic marks used was doubled. The inclusion of the three individual tests showing the lower correlations raised the correlations for the group of tests but very little. *The importance of the first fact is not to be overlooked,* for it urgently suggests that the low correlations found between *the Army Test scores and the averages of the marks for* either semester of the freshman year, and especially the averages of the marks for the first quarter of the sophomore year, were in no small measure due to errors of judgment in the estimations of the achievements of students. Furthermore, when the scores of the two Army Tests were combined the correlations between the combined scores and the averages of the marks for the two semesters was +.47. Also when the scores of the ten selected tests of the two Army Tests were combined into a single score, the correlation with the averages of the marks for the two semesters was likewise +.47. *Thus it is clear that to obtain high correlations between* the scores made in mental tests and averages of academic marks, it is essential to have more accurate measures of the academic achievements of students than the subjective estimates of the instructors. This of course does not imply that it is either unnecessary or not worth while to increase the reliability and the scope of the tests. It does suggest, however, that because of the inaccuracies of college marks as measures of academic achievements it will not be feasible to accurately determine the full prognostic value of the tests.

That academic marks are not at all accurate indices of achievement or of abilities acquired is not a sheer guess nor even just a matter of opinion. The composition ability of some ninety-eight sophomore students was measured by having them write two half-hour themes and then having the themes rated for general merit by three college instructors of English with the use of the Thorndike Extension of the Hillegas Scale for Measuring English Composition. The correlations between the rhetoric marks of the freshman year and the quality of the themes were as follows:

First semester rhetoric marks and qualities of composition A r = +.45
First semester rhetoric marks " " " " B r = +.42
Second semester rhetoric marks " " " " A r = +.41
Second semester rhetoric marks " " " " B r = +.39

The correlation between the rhetoric grades for the two semesters was +.68, while that between the two themes was +.49. Making corrections for attenuation, the correlation between the ability to make marks in freshman rhetoric and the ability to write English compositions was only +.72. This is only one measurement of the relationship in question and hence is not to be taken too seriously. There is no reason to suppose that rhetoric marks are less accurate than the marks in other college subjects. With no higher correlation, however, than +.72 between the ability to make academic marks and the ability to achieve results, one cannot expect very high correlations between scores in mental tests and academic marks, even though the tests have a high value for prognosticating probable academic achievement. This condition of course places upon the psychologist working to derive more adequate mental tests for college use an additional feature of uncertainty to cope with.

The most important finding in connection with the use of the Army Tests at the University of Minnesota, however, was the revelation of distinct group differences among the several college student bodies. In four of the college groups to which the tests were given in the two successive years the personnel of the groups probably changed but little from the one year to the next. A comparison of the median achievements of these groups afforded some suggestive results. In comparing the positions of the medians of the eight individual tests of the two Army Tests in the case of the men and the women in the College of Science, Literature and the Arts the median men stood higher than the median women in five of the tests of Form E. They also stood higher in four out of the five corresponding tests of Form 6. In three of the tests of Form E the median men stood lower. In each of the three corresponding tests of Form 6 they also stood lower. Upon making a similar comparison in the case of the men in the College of Science, Literature and the Arts and the men in the College of Engineering, the median men in the College of Engineering stood higher in five of the tests of Form E. In four of the five corresponding tests of Form 6 they also stood higher. At the same time the median men in the College of Engineering stood lower in the other three tests. In two of the three corresponding tests of Form 6 they also stood lower. This lends support to the presumption that there is probably a much closer correspondence between the relative sizes of the medians of two corresponding groups of individuals in two similar tests than between the relative sizes of the medians for two different groups in the same or similar series of tests. Assuming this to hold true for the large group of male students in the College of Science, Literature, and the Arts, the medians of this group were taken as standards in all the tests and the per cent of each of several groups doing as well as or better than this standard was determined for each test in the two groups of tests, Form E and Form 6. In so far as the assumption does not hold true, obviously, the effect would be to lower the correspondence where the corresponding groups took two different series of similar tests.

After the per cents of overlapping for each test for each of the available groups—the College of Engineering students and the College of Dentistry students—had been ascertained, the tests in each group of tests were arranged for each college group according to the amounts of overlapping. When the correlations were then determined between the various arrangements it was found that the coefficients for each two different college groups, derived by Pearson's shorter method, were +.07, —.04, +.39, and —.10 respectively, but that when the relationships between the two arrangements for the two sets of tests for the corresponding groups of the same college were determined the coefficients were +.71 for corresponding groups of one college and +.68 for the corresponding groups of the other college. This means that the second freshman class tested in the College of Engineering stood high in the same tests as did the first freshman class, also low in the same tests as did the first freshman class tested. The same condition was found to hold true for the two freshman classes tested in the College of Dentistry. On the other hand, the freshman classes in the College of Dentistry did not stand highest in those tests in which the freshman classes of the College of Engineering did, nor did they stand lowest in those tests in which the freshman classes of the College of Engineering stood lowest. In fact, there was practically no similarity or relationship evident when the freshman classes of the different colleges were compared with one another as is shown by the average of the four correlations above; namely, $r = +.08 \pm .05$.

These indications strongly suggest that in so far as students' interests tend on the whole to lead them to seek the occupations for which they are the better fitted by nature, mental tests may in time be used not only as a means of *predicting* a student's probable chances of success in college work in general but also as a basis of predicting his chances of success in doing the work of each of the various technical colleges and at the same time as a basis for giving the student more accurate advice in the matter of selecting a vocation than can now be given. This task will be long and arduous, and one demanding the fullest degree of co-operation on the part of educational leaders, especially those in the universities, but the results will surely repay for the efforts many-fold.

PROPHESYING ARMY PROMOTION

By S. C. Kohs, Psychologist to Court of Domestic Relations, Portland, Oregon, and K. W. Irle, Reed College, Portland, Oregon

The human mind has manifested from time immemorial an insatiate hunger for those phenomena and traits which could answer the question, no matter how indirectly and inadequately, "What may I expect on the morrow,—next month,—next year?" Our star-gazers, palm-readers and our bump-feelers are still with us in spite of persistent efforts on the part of scientific criticism to discredit them, and the intense eagerness to know that which the future may hold in store for us has increased in the last few decades to immeasurable proportions.

It is not surprising that many have turned with hope toward psychology as the analyzer of individuality, of traits, of customs, of events; toward psychology as a synthesizer of more or less fragmentary mental evidence for the purpose of more clearly indicating or prognosticating what one might expect in the future, granting that a given set of causes would inevitably be active.

This tendency toward developing a prophetic procedure is taking more concrete and definite form in connection with intelligence testing than with any other phase of experimental psychology. Thus we note one study on "The Relation of Mental Testing to School Administration, with Special Reference to Children Entering School,"[1] in which one of the objects was "to offer predictions as to the probable advancement of each child through the primary grades,—this prediction being based primarily on the potential mental capacity as shown by the Binet mental test." (p. 4). A similar study is now under way at Stanford University by W. M. Proctor, dealing with the prediction of progress of students through the secondary school. Terman's book "The Intelligence of School

---

[1] By Virgil E. Dickson, in *Normal Seminar Bulletin A*, No. 1, July, 1917, Dept. of Educ. State Normal School, Cheney, Wash. See also L. M. Terman: *The Intelligence of School Children*, Houghton Mifflin, 1919.

Children" is a splendid contribution to this general subject. The results of these investigations lead one to a very optimistic outlook regarding the possibilities of prophesying school progress. The psychological work in the army was largely concerned with predicting whether a raw recruit showed promise of being an asset or a liability to the army organization, and if an asset whether as a leader or a follower. In its pamphlet[2] the Psychological Division of the Army presents ample proof that the Army Intelligence Tests have been an efficient prophesying agency. And, finally, Terman is accumulating data which seem to point toward a constancy of relationship between life age and mental age, which if it is proven to exist will materially assist the prophecy of success or failure of an individual in various lines of activity. The increase of interest and research in trade tests and special aptitude tests will add data of inestimable value making for greater efficiency in selecting promising material for special lines of endeavor.

The present study is an attempt to determine to what extent Reed College could have predicted the progress of 116 of its students who entered the service of the army or the navy. The data upon which such predictions might have been based would have been (a) the quality of their college work, and (b) faculty estimates regarding (1) their physical qualities, (2) their intelligence, (3) their leadership, (4) their personal qualities, and (5) their general value to the service.

The Original Data

The material for investigation came from three distinct sources.

Source A. RATINGS. Three judges, members of the Reed faculty, rated each of the 116 students by means of Scott's Rating Scale.[3] Information was thus obtained regarding the faculty's estimate of each of these men on physical qualities (3 points the lowest, to 15 points highest), intelligence (3 to 15), leadership (3 to 15), personal qualities (3 to 15), and general value to the service (8 points the lowest, to 40 points the highest).

Source B. MARKS. The college marks were obtained from the Registrar's office and were grouped under three headings, as follows:

---

[2] See *Army Mental Tests: Methods, Typical Results and Practical Applications*, Nov. 22, 1918, p. 23. Wash., D. C.; L. M. Terman: "The Use of Intelligence Tests in the Army," *Psychol. Bull.* 1918, 15, pp. 177-187; "The Measurement and Utilization of Brain Power in the Army," *Science*, 1919, 49, pp. 221-226, 251-259.

[3] War Dept. Adjutant General's Office. Forms CCP-1102, 5-18-18 and CCP-1104, 5-22-18. See also "The Rating Scale," *Psychol. Bull.* 1918, 15, pp. 203-206.

"Natural Science" included marks in biology, chemistry, mathematics, physics, psychology, astronomy, geology and natural science.

"Social Science" included marks in economics, education, history, politics, political science, sociology, philosophy and social ethics, and

"Languages" included marks in English, Greek, Latin, Romance Languages, Germanic Languages, and for want of a better place to insert them, art, surveying and mechanical drawing.

The writers were aware that these groupings were artificial and arbitrary, but some grouping was necessary and these are perhaps as free from serious objections as any.

The following table presents an analysis of the number of marks in each subject-group possessed by various frequencies of students; thus, 7 students had only one mark in natural science, 13 students had only one mark in social science, 16 students had 4 marks in the language and fine arts subjects, etc.:

	NO. OF STUDENTS HAVING MARKS IN			
No. of Marks	*Nat.* Sci.	*Soc.* Sci.	*Lang.* and F. Arts	*Total*
1	7	13	6	26
2	25	9	22	56
3	7	9	2	18
4	14	7	16	37
5	4	11	5	20
6	9	11	8	28
7	4	7	8	19
8	2	5	10	17
9	6	3	7	16
10	3	6	10	19
11	2	1	2	5
12	0	5	3	8
13	2	0	2	4
14	1	2	0	3
15	5	4	1	10
16	2	2	1	5
17	4	3	2	9
18	0	1	1	2
19	1	1	1	3
20	1	1	0	2
21	2	0	1	3
22	0	0	0	0
23	0	1	0	1
24	0	0	0	0
25	0	1	0	1
Total	101	103	108	312

Source C. ARMY RANK. The college attempted to keep a careful record, brought up to date, of the progress which each of its students was making both in the army and in the navy. The ranks were those which each held the day the armistice was declared. The ranks ranged from Private to Major. The following table and graph present the frequency of each rank.

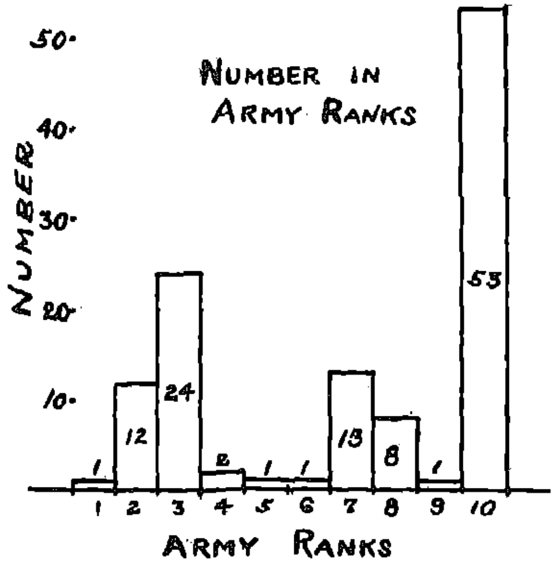


TABLE SHOWING THE FREQUENCY OF EACH MILITARY RANK

Rank	No.	Rank	No.
1. Major	1	6. Bat. Sergt.-Major	1
2. 1st. Lieut.	12	7. Sergt.	13
3. 2nd. Lieut. (Ensign)	24	8. Corporal	8
4. Reg. Sergt.-Major	2	9. 1st. Class Private	1
5. 1st. Sergt.	1	10. Private	53
		Total	116

Statistical Treatment

The statistical treatment divided itself into seven portions: (1) the inter-correlations between the estimates of the various judges, (2) the correlations between the consensus of opinion of the judges regarding the five traits and final army rank, (3) the correlations between the five traits, (4) the correlations between the five traits and the marks attained in the various subject-groups, (5) the average marks attained in the different subject-groups by each of the ten army ranks, (6) the correlations between final army rank and marks obtained in the three school subject-groups, and (7) the average ratings in each of the five traits received by each of the ten army ranks. It would have been valuable to have obtained an estimate of the reliability of each judge, but it would have involved time and labor which we could not very easily request. However, from similar ratings in the army and from past information regarding the reliability of judges in estimates of this nature, we may take for granted a rather high degree of reliability of judgment. This is evident if we note the intercorrelation between the ratings of the various judges, especially between judges 2 and 3.

In view of the fact that the length of service varied from 4 months to 18 months it was felt necessary to divide the 116 students into two groups, one consisting of short-service men and the other of long-service men. In the latter group were placed all who had been in the army or navy more than 10 months. This division was important because the long-service men would presumably have had a better opportunity to obtain promotion. In fact, the real test of the validity of various prognosticating criteria will depend upon how accurately these criteria have "sized-up" this group of long-service men. The short-service men number 26, the long-service men 90.

The correlation-coefficients mentioned herein were all obtained by means of the formula:

$$r = \frac{\Sigma fxy}{\sqrt{\Sigma fx^2} \cdot \sqrt{\Sigma fy^2}}$$

modified by one of the writers (S. C. K.) to correct some factors which make for too great a deviation from the true Pearson "r-formula:"[4]

$$r = \frac{\Sigma xy}{n\sigma_1\sigma_2}$$

Results

(1) *The correlations between the estimates of the various judges.*

(a) Judge 1 and Judge 2.

Judge 1 was not as reliable as either 2 or 3, owing to his being somewhat pressed for time in connection with war-service matters. The correlations between the estimates for the five traits were: (Coefficients are positive unless otherwise indicated), (the probable errors of each coefficient are given in the summary table at the end of this article):

---

[4] Discussed more fully by one of the writers (S. C. K.) in his forthcoming monograph on the "Block Design Test."

(1) physical qualities	.34
(2) intelligence	.40
(3) leadership	.37
(4) personal qualities	.39
(5) value to the service	.34

(b) Judge 1 and Judge 3.

The correlations between the estimates of these judges were somewhat better, though still unusually low:

(1) physical qualities	.53
(2) intelligence	.29
(3) leadership	.50
(4) personal qualities	.38
(5) value to the service	.30

(c) Judge 2 and Judge 3.

The estimates of these two judges may be regarded as quite reliable. The correlations between their judgments follow:

(1) physical qualities	.41
(2) intelligence	.51
(3) leadership	.65
(4) personal qualities	.62
(5) value to the service	.72

The estimates of physical qualities showed the lowest correlation, whereas those of value to the service showed the highest correlation.

The results of this analysis of the ratings of the three judges showed rather mediocre reliability of capacity to judge a student's rank with regard to any one of these traits. Of course, one of the judges may have been quite expert, but the bare data themselves do not reveal such expertness. In fact, an analysis of the correlations between final rank obtained in the army and the average trait estimates of the three judges, in the next division of our study, brings into still greater question this subjective method of prognosticating progress in the Army. Averaging the correlations of the estimates of each of the three judges for the five traits we obtain the following:

(2) intelligence	.40
(1) physical qualities	.43
(5) value to the service	.45
(4) personal qualities	.46
(3) leadership	.51

It is interesting to note that the judges showed least agreement on " intelligence " and most on " leadership."

(2) *Correlations between the consensus of opinion of the judges regarding the five traits and final army rank.*

(a) Short-service men.

The correlations between the ranks which these men had attained during their short-lived stay in the army and the estimates of the judges of their capacities in each of the five traits enumerated, are as follows:

(1) physical qualities	.21
(2) intelligence	.39
(3) leadership	.08
(4) personal qualities	.23
(5) value to the service	.30

Of course, one might argue that these results clearly indicate that because these men were in the army but a short period, for this reason these correlations are low. However, the correlations are just as low for the long-service men.

(b) Long-service men.

(1) physical qualities	.28
(2) intelligence	.14
(3) leadership	.14
(4) personal qualities	.21
(5) value to the service	.33

Although estimates of "value to the service" and final army rank show the highest or next to the highest correlations in both groups, this is *not* true for estimates of " intelligence,"—which is *highest* in its correlation with final army rank in the *short-service* group, and *lowest* in its correlation with final army rank in the *long-service* group. It cannot be said that " intelligence " becomes less of a factor for success the longer one remains in the army. But here again one questions the value of *subjective* estimates of amounts of trait manifested by various individuals. However, there is the other alternative: one might maintain that our judges *are* expert and that the explanation for these low correlations is the haphazard fashion in which the army attacks the problem of advancing its men or that it considers other factors, not here enumerated, the important ones for success.

(3) *The correlations between the five traits.*

The estimates of all the three judges were massed for the purpose of this inquiry.

One looks forward with the greatest anticipation to that study which will attempt to analyze, psychologically and statistically, the factors entering into the moulding of human judgments. What are the influences, conscious or unconscious, which make for idiosyncrasy, what is the mechanism of judgment, what makes for constancy, for reliability, how easily are judgments changed, what are the subjective criteria upon which judgments are based? The present portion of our investigation is somewhat apart from our main thesis. Yet there is one striking feature of the data which may be suggestive to someone interested in this branch of research. There has been a good deal written regarding the Spearman-Hart-Burt explanation of intelligence as a "general common factor." Evidence to that effect is deduced by a statistical analysis of the results of various psychological tests each of which measures some amount of this "general common factor." The "Hierarchy of Coefficients" lends added weight to this hypothesis. We may here draw an analogy. It seems very probable that when one is passing a subjective judgment on the question of whether Person A possesses a certain amount of trait *a*, or *b*, or *c*, or *d*, that his judgment of practically all of these is affected by some constant factor *x*. For example, here is Tom Jones. Bill Smith is requested to record a personal estimate of his character, habits, self-control, intelligence, sociability,—whether excellent, good, fair, very poor. What probably occurs when Bill estimates, is that each of his judgments is affected by a constant factor, possibly unconscious, such as "Tom Jones is an excellent fellow. I like him because his ideas are very attractive to me." This example is not typical, of course, but is merely utilized to illustrate the point. That some such condition possibly exists is suspected from this table of data, which shows, somewhat, a hierarchy of coefficients.

	1. Physical Qualities	2. Intelligence	3. Leadership	4. Personal Qualities	5. Value to the Service
1. Physical Qualities..........		.27	.54	.45	.63
2. Intelligence................	.27		.52	.56	.62
3. Leadership..................	.54	.52		.60	.71
4. Personal Qualities..........	.45	.56	.60		.80
5. Value to Service...........	.63	.62	.71	.80	

It is of interest to note (a) that the smallest correlation is between physical qualities and intelligence (.27), and the highest between personal qualities and value to the service (.80), (b) that physical qualities and intelligence show the least correlations with the other traits whereas value to the service shows throughout higher correlations.

(4) *The correlations between each of the five traits and the marks attained in the various subject groups.*

The question, "To what degree do school marks and the trait estimates of judges coincide when prognosticating army success," is answered by the following correlations:

	Physical qualities	Intelligence	Leadership	Personal qualities	Value to the Service
Marks in Nat. Sci.	.15	.77	.40	.40	.38
Marks in Soc. Sci.	.34	.57	.41	.44	.46
Marks in Lang. and Fine Arts	.30	.66	.36	.33	.38

The estimates of the three judges were averaged (Arith. Mean) in determining these correlations. It may be of interest to note that physical qualities showed the least correspondence with school marks, whereas intelligence showed the highest. In other words: Of all the five enumerated traits, intelligence was the most important in determining success in school work, whereas physical qualities, although important, apparently were the least important of all. There is that much to the credit of the faculty and the grading system!

There may be some who would be interested in a further analysis: It seems from these data that intelligence is more of a factor for success in the natural sciences, *least* in the social sciences, with languages occupying a middle zone!

It is evident from the above table that with the exception of intelligence, school marks inadequately coincide, if at all, with the judges' ranking of these men in order of their abilities. The two criteria, apparently, are separate,—and, apart from intelligence, measure a somewhat different array of characteristics.

(5) *The average marks attained in the various subject groups by each of the ten army ranks.*

Since the short service men had representatives in only three ranks, private, sergeant, and second lieutenant, and because of the difference in service time, the short and long-service men will be considered separately. The following tables present the number of men of each rank in the two groups:

SHORT SERVICE MEN

Rank	No. of Men
3. 2nd. Lieut. (Ensign)	4
7. Sergeant	2
10. Private	20
Total	26

LONG SERVICE MEN

Rank	No. of Men
1. Major	1
2. 1st. Lieut	12
3. 2nd. Lieut. (Ensign)	20
4. Reg. Sergt.-Major	2
5. 1st Sergt.	1
6. Bat. Sergt.-Major	1
7. Sergt.	11
8. Corporal	8
9. 1st Class Private	1
10. Private	33
Total	90

In considering the long service men, only ranks (2), (3), (7) and (10) will be compared since the number of men in the other ranks number no more than one or two.

(a) Short-service men.

Before presenting an analysis of the marks attained, a few words might be said in preface regarding the Reed College marking system (from the current Reed College Catalogue):

"Grades in courses of study are awarded on a scientific rather than a personal basis, with definite credit for quality as well as for quantity of work. Until all school work can be measured by scales, made up of units that are equal in a defined sense, the best available grading is one of relative position in a series. The nearest approach to such a scientific basis for awarding college credits appears to be a distribution following the normal probability curve, skewed to take account of the effect of selecting the student body.

"Reed College has, from the outset, used ten grades, whose definitions have such a scientific basis.

Grades.............	1	2	3	4	5	6	7	8	9	10
Proportion of Students.........	....	5%	10%	15%	20%	25%	15%	6%	4%	....

"Grades 1-5 indicate that a student stands in the upper half of an average class; grades 6-10 indicate that he is in the lower half. For example, 2 designates the work which will be done (in the long run) by the best 5% of all students, and 6 the work done by that quarter of an average class standing just below the middle.

"Grade 1 is rarely given, representing a degree of excellence attainable by not more than one student in four or five hundred; grade 10 records correspondingly bad failures. The lowest passable grade is 8; 9 is for ordinary cases of failure. The grades cannot be interpreted in qualitative terms, as good, poor, A, C, 90%."

In the following table are presented the average school marks of the three different ranks:

AVERAGE MARKS IN

Rank	Nat. Sci.	Soc. Sci.	Language and Fine Arts
3	4.8	5.9	5.2
7	5.7	6.2	6.1
10	5.3	5.0	5.8

Rank (7) (Sergeants) had, on the whole, poorest marks throughout. Rank (3) (2nd Lieut. Ensign) was clearly superior to (10) (Private) on the basis of marks as one would naturally expect, this being especially true for marks in natural science. It may or may not be surprising that the evidence was reversed for marks in social science, rank (3) being inferior to (10). Summarizing the data of the table it may be said that those of higher rank obtained higher marks in the natural sciences and languages than those of lower rank, the situation being reversed for proficiency in the social sciences.

(b) Long-service men.

In the following table are presented the average school marks of the four different ranks:

AVERAGE MARKS IN

Rank	Nat. Sci.	Soc. Sci.	Language and Fine Arts
2	5.3	4.9	5.6
3	6.1	5.5	6.3
7	5.1	5.4	6.1
10	6.1	6.0	5.9

Rank (2) is, throughout, superior to rank (10) especially in the social sciences. And as a medium of prognostication these subjects would seem to have the advantage over other subjects in the college curriculum, this being the direct opposite of the indications apparent in the case of the short-service men in the previous division. Rank (7) was, on the whole, superior in marks for each group over rank (3). A matter worth mentioning is the variability of marks within each rank. A glance through the various school marks which the members of the different groups had obtained reveals as great an average variability of marks within each rank as is evident between the averages of the different ranks. It is questionable whether, except in extreme cases, one can rely on a school mark in any subject as an aid in prognosticating army promotion. This matter is more emphatically brought to our attention in the next item to be considered.

(6) *The correlations between final army rank and marks obtained in the three school subject groups.*

The correlations for the short and the long-service men are presented separately.

SHORT-SERVICE MEN

Correlation Between Final Rank and Marks in	Co-efficient
Nat. Science	.12
Soc. Science	.12
Language and Fine Arts	.06

These correlation coefficients are so small that little if any diagnostic significance can be attached to the school marks for this group of men.

LONG-SERVICE MEN

Correlation Between Final Rank and Marks in	Co-efficient
Nat. Science	.14
Soc. Science	.24
Language and Fine Arts	.01

The coefficients here are also small, but as in the previous section, marks in social science for the long-service men seem more reliable criteria for prognosticating army promotion, but even then, their value is greatly limited because of the many exceptions.

On the whole, school marks, although they might aid prognostication, cannot be depended upon for sole support in this effort. In fact, poor as judges' estimates may be, they seem somewhat superior to school marks for diagnosing army progress. This will be more apparent from the data in the succeeding section.

(7) The average ratings in each of the five traits received by each of the ten army ranks.

SHORT-SERVICE MEN

Rank	Physical Qualities	Intelligence	Leadership	Personal Qualities	Value to Service	Total
3	12	14.3	10.3	13.5	33	83.10
7	13.5	10	11.5	11.5	28	74.50
10	11	11.2	10.6	11.5	27.8	72.10

Although the separate traits with the exception perhaps of "value to the service" do not demonstrate any clear correlation between height in rank and height in trait-score, nevertheless the totals do show this clearly. Comparing the highest rank (3), with the lowest, (10), the former throughout shows higher trait-scores than the latter with the exception of "leadership." "Intelligence" and "Value to the Service" show the greatest difference in trait-score between the highest and lowest ranks. This helps confirm the earlier claims for these two traits as more efficient for prognostication than any of the other three.

LONG-SERVICE MEN

Rank	Physical Qualities	Intelligence	Leadership	Personal Qualities	Value to Service	Total
2	11.9	12.5	12.2	12.9	34.8	84.30
3	12	10.4	10.7	11.5	29.7	74.30
7	11.7	11.7	11.3	12.2	29.7	76.60
10	10.4	10.1	10.3	10.9	26.2	67.90

Here again rank 2, as well as rank 3, in each of the five traits, demonstrated considerable superiority over rank 10; especially is this true for rank 2. Again, with the exception of "value to the service" the separate traits do not show a clearcut correlation between height in rank and superiority in trait-score. It is of interest to note that rank 2, averaged more than 16 points higher in total score than rank 10. As was the case with school marks, so here, with regard to judges' estimates, rank 7 is apparently superior to rank 3 in the opinion of the judges, and apparently the army has not utilized the same criteria in its judgment and placing of these men.

SUMMARY OF THE CORRELATIONS

Cor. No.	Cases	Facts Compared	Cor.	P.E.
1	113	Estimates of Judge I and II re Physical Qualities	.34	.06
2	113	" " " " " Intelligence	.40	.06
3	113	" " " " " Leadership	.37	.06
4	113	" " " " " Personal Qualities	.39	.06
5	113	" " " " " Val. to the Service	.34	.06
6	96	" " " " III Physical Qualities	.53	.05
7	76	" " " " " Intelligence	.29	.07
8	39	" " " " " Leadership	.50	.08
9	60	" " " " " Personal Qualities	.38	.07
10	35	" " " " " Val. to the Service	.30	.10
11	96	" " II " " Physical Qualities	.41	.06
12	76	" " " " " Intelligence	.51	.06
13	39	" " " " " Leadership	.65	.06
14	60	" " " " " Personal Qualities	.62	.06
15	35	" " " " " Val. to the Service	.72	.06
16	26	Est. of Judges for Phys. Qual. and Rank 1-9 mos.	.21	.13
17	26	" " Intellig'nce " " " " →	.39	.12
18	26	" " Leadership " " " "	.08	.13
19	26	" " Pers. Qual " " " "	.23	.12
20	26	" " Val. to Ser. " " " " →	.30	.13
21	90	" " Phys. Qual. " " 10-18 "	.28	.06
22	90	" " Intellig'nce " " " "	.14	.07
23	90	" " Leadership " " " "	.14	.07
24	90	" " Pers. Qual. " " " "	.21	.06
25	90	" " Val. to Ser. " " " " →	.33	.06
26	21	Nat. Sci. marks " " 1-9 "	.12	.15
27	23	Soc. Sci. " " " " "	.12	.15
28	25	Lang. F.A. " " " " "	.06	.13
29	80	Nat. Sci. 10-18 "	.14	.08
30	80	Soc. Sci. " " " " "	.24	.07
31	85	Lang. F.A. " " " " "	.01	.08
32	101	Nat. Sci. " and Judges' est. Phys. Qual.	.15	.07
33	102	Soc. Sci. " " " " " "	.34	.06
34	108	Lang. F.A. " " " " " "	.30	.06
35	100	Nat. Sci. " " " " Intelligence	.77	.03
36	103	Soc. Sci. " " " " "	.57	.04
37	108	Lang. F.A. " " " " "	.66	.04
38	101	Nat. Sci. " " " " Leadership	.40	.06
39	103	Soc. Sci. " " " " "	.41	.06
40	108	Lang. F.A. " " " " "	.36	.06
41	101	Nat. Sci. " " " " Personal Qual.	.40	.06
42	103	Soc. Sci. " " " " " "	.44	.06
43	107	Lang. F.A. " " " " " "	.33	.06
44	100	Nat. Sci. " " " " Val. to the Ser.	.38	.06
45	102	Soc. Sci. " " " " " " "	.46	.05
46	108	Lang. F.A. " " " " " " "	.38	.06
47	116	Est. of 3 Judges as to Phys. Qual. and Intell'nce	.27	.06
48	116	" " " " " " Leadership	.54	.05
49	116	" " " " " " Pers. Qual.	.45	.05
50	116	" " " " " " V. to Ser.	.63	.04

SUMMARY OF THE CORRELATIONS—Continued

51	116	Est. of 3 Judges	Intelli'ence	and	Leader	.52	.05
52	116	" " "	"	"	Pers. Qual.	.56	.04
53	116	" " "	"	"	V. to Ser.	.62	.04
54	116	" " "	Leadership	"	Pers. Qual.	.60	.04
55	116	" " "	"	"	V. to Ser.	.71	.04
56	116	" " "	Pers. Qual.	"	V. to Ser.	.80	.03

Conclusions.

Summarizing the results of our analysis the following would be our conclusions:

(1) School marks are rather inefficient instruments for determining whether a student will make good progress in the army.

(2) Human judges, with all their frailties, are, on the whole more efficient prognosticators of progress than the school marks which students obtain.

(3) Of all the criteria for prophesying success, the safest *are, first, judges' estimates of value to the service, and second,* judges' estimates of intelligence. In both cases, however, the correlations are low. If objective, rather than subjective, estimates of these traits had been used, the correlations might have been increased considerably. We already possess objective instruments for measuring "intelligence," why not produce similar instruments for measuring "value to the service?"

(4) We have taken for granted, throughout this study, that the army was correct in its selections. We recognize that this *assumption is not wholly valid; the real fault may not lie at* all with our criteria but rather with the current methods in the army by means of which men are selected for superior positions.

(5) The low correlations should be therefore explained as being due in part to

(a) Imperfections in our standards of judging individual efficiency, marks and the estimates of judges;

(b) Imperfections in the system of army promotion;

(c) *Differences in the factors upon which the army bases* its promotion and those upon which school marks and judges' estimates are based.

(6) Although school marks and judges' estimates may serve as aids in prophesying army progress, other criteria must be relied upon for any satisfactory development of a prognosticating machinery.

THE DEGREE OF PH.D. AND CLINICAL PSYCHOLOGY

By EDGAR A. DOLL, Psychologist, N. J. State Dept. Institutions and Agencies

There is reason to protest against the use of the degree of Ph.D. as a *sine qua non* in the "certification" of a clinical psychologist. It may have occurred to some who have enjoyed the discussions in this JOURNAL of what constitutes clinical psychological expertness that one certifies himself as a clinical psychologist by the nature and character of his work and the consequent reputation therefrom ensuing.

There seems to be a general agreement, however, that the degree of Ph.D. is an essential. The argument is presumably based on two considerations (a) that the degree is a testimony of advanced work and scholarly accomplishment, and (b) that it is something of a guarantee of superior general mental ability. Against (a) it may be protested that the attainment of this degree in academic psychology is no guarantee of either information or ability in clinical psychology, although undoubtedly it is a desirable and perhaps necessary basis for clinical psychology. Against (b) it may be argued that while those who hold the Ph.D. may be willing to admit its selective influence as a measure of general mental ability there will be others who might contend that clinical psychology demands specific technical ability which may be more or less independent of general ability at the Ph.D. level of general ability.

It is one of the functions of a clinical psychologist to discover the exceptional case. He specializes in individual differences. The diagnosis of ability is determined without error only when the standard employed excludes all in whom the trait is absent and also includes all in whom it is present. Applying clinical methods to the diagnosis of clinical psychological ability, can it be said that the Ph.D. degree is valid as a fixed condition of clinical psychological ability? We have seen in the Army that such was not the case. The overlapping of men with the degree who were failures as clinical psychologists and men without the degree who were successes was certainly too great to permit the use of the attainment of this degree as anything but a general criterion of clinical psychological ability.

Some reasons for this are fairly obvious. Ability in the field of individual mental diagnosis is very largely a matter of *specific* training rather than general preparation. In the long run the man with a Ph.D. degree has the advantage in ease and rapidity of assimilating such training, but his attainment in general psychology is no guarantee of the specific preparation required. Moreover, a man of fair general intelligence without the degree is able under good instruction to gain this specific ability independently of the Ph.D. degree (which is after all only certification for *formal* work *in residence* at a college of accredited standing).

Clinical ability is also founded on experience with clinical material under clinical conditions. This ability cannot be imparted *ex cathedra* in the classroom. Hence a man with the Ph.D. degree who has specialized in the theoretical or academic considerations of clinical psychology is not qualified until he has served his "interneship," which is perhaps equally as important as the Ph.D. degree itself.

Moreover mental diagnosis is nearly as much an art as it is a science. A successful clinical psychologist must have a successful "clinical personality," the "clinical temperament," the ability to obtain and maintain that *rapport* which we so frequently hear mentioned.

We therefore maintain the following general propositions:

1. Clinical psychological ability demands specific training in the several allied fields of mental diagnosis (such as physiology, psychiatry, anthropometry and education, for example) as well as general training in academic psychology.
2. This ability is based on experience and specific training as well as on formal academic preparation.
3. This experience and specific training may be obtained independently of the Ph.D. degree.
4. Some men with the Ph.D. degree are failures as clinical psychologists and some men without the degree are successes. The extent of overlapping is not inconsiderable.
5. Therefore, while the degree of Ph.D. may be a desirable adjunct to clinical psychological ability, it is not a necessary prerequisite.

In conclusion it is well to emphasize that we do not favor dispensing with the Ph.D. as *a* qualification for certification.

Undoubtedly many Doctors of Philosophy have become psychological clinicians, just as many clinicians have become Doctors of Philosophy. But it is important to emphasize that a candidate for certification ought to be free to offer demonstrated ability or knowledge in clinical psychology *in lieu of* the Ph.D. degree. The American Psychological Association does not insist on the Ph.D. degree as an unconditional qualification for membership; a candidate may offer the equivalent of the degree in terms of demonstrated ability to pursue work of high character in psychology. There seems neither need nor justification for greater rigor of qualifications in the subordinate field.

MINOR STUDIES FROM THE PSYCHOLOGICAL LABORATORY OF INDIANA UNIVERSITY

VI. THE INFLUENCE OF (a) INADEQUATE SCHOOLING AND (b) POOR ENVIRONMENT UPON RESULTS WITH TESTS OF INTELLIGENCE[1]

By LUELLA WINIFRED PRESSEY.

I. *Problem.* The present paper is essentially a continuation of two studies which have already been briefly reported in this JOURNAL; the first dealt with the comparative intelligence of country children and city children, and the second compared in a similar fashion children from good and from poor homes.[2] In these studies two group scales applicable from the third grade through high school were used. In each instance the test findings were looked to for aid in an understanding of various sociological and economic facts also discovered, in the course of the survey, with regard to the districts and families studied. And in each case the tests showed marked differences; the city children rated distinctly above the country children, and children whose fathers were day laborers were found strikingly below the children of professional men in "native endowment."

In the course of the survey, however, many facts were found which seriously brought in question the validity of these findings. Thus most of the country schools were "six months schools" and many of the children failed to attend regularly even during this brief period when the schools were open. The country district studied is, in fact, in a county notorious in the state for the inefficiency of its school system; the teachers are incapable and ill-trained, and the equipment miserably inadequate—in some instances the children were entirely without paper and pencils, the school work being done altogether on slates or on the blackboard! The group scale of intelligence given these country children presupposed (as do all such scales for use in the upper grades) a considerable degree of literacy and a fair reading vocabulary, and it involved the use of pencil and paper. It seemed, therefore, very possible that the poor showing made by these children might be the result quite as much of their inadequate schooling as of any lack in native intelligence. Somewhat analogous difficulties were encountered in evaluating the data obtained from children from homes at different economic levels. In good homes there is a background of general information and culture, and a wealth of reading, which (it would seem) cannot but facilitate work on such tests. In the homes of the quarry hands and factory workers of the city studied, such influences are notable for their absence; in fact, the parents are often illiterate. In short, these various accessory factors were felt to be so important, and so pervasive in their influence, that in the papers presenting these two studies it has been urged that results with a scale of performance tests (for the country children), and with tests for measuring home culture (as a check upon the "occupation of parents" data) would be necessary, before the findings with the tests of intelligence could be satisfactorily interpreted.

---

[1] The paper was presented, in slightly different form, at the meetings of the American Psychological Association, Cambridge, Mass., Dec., 1919.

[2] Pressey, S. L., and Thomas, J. B., "A Study of Country Children in (1) A Good and (2) A Poor Farming District by Means of a Group Scale of Intelligence," *Journal of Applied Psychology*, Vol. 3, pp. 283-286, 1919, and Pressey, S. L. and Ralston, Ruth, "The Relation of Occupation to Intelligence as It Appears in the School Children of a Community," *Journal of Applied Psychology*, Dec., 1919, pp. 366-373.

It has been found possible, however, to get back of such special influences for the most part, by the simple expedient of examining the children from these different groups at their entrance to school, or soon after, with the "Primer Scale."[3] This brief scale of intelligence does not involve literacy, nor school training; children from country and city should thus meet the examination on equal terms. And the scale is given to the children at so early an age that the home influences just referred to might be expected to have operated to a much less extent than later; particularly is it important that the children are tested before they have learned to read readily—any effects coming from the superior opportunities for reading in the good home are thus largely avoided. Comparison of results obtained from the younger children, using the Primer scale, with results from the older children, using the more usual type of examination, should in fact not only make clearer the relative mental ability of the groups studied; the comparisons should aid in evaluating the scales involved, and also in estimating the general importance of the various environmental factors mentioned, in mental test work.

---

[3] For a description of the scales mentioned in this paper see Pressey, S. L. and L. W., A Group Point Scale for Measuring General Intelligence, *Journal of Applied Psychology*, Vol. II, 1918, pp. 250-269. "Cross-out" Tests, *Journal of Applied Psychology*, Vol. III, 1919, pp. 138-150; or Pressey, L. W., A Brief Group Scale of Intelligence for Use in the First Three Grades, *J. Ed. Psychol.*, Sept., 1919. The material is also briefly discussed in the *Bulletin of the Extension Division, Indiana University*, Vol. V, No. 1, 1919.

II. *Results.* The first data obtained for the present study with the Primer scale were from 183 country children six, seven, and eight years of age—all the children of these ages in fourteen country schools which had already been surveyed the year before with a scale for use in the upper grades (the "Schedule D" or "Group Point Scale") containing ten tests of memory, controlled association, arithmetical reasoning, and so on. These results obtained the year before with children 10-14 years old showed only 20% of the country children rating above the median for their age, when compared with norms obtained from city children. If, now, this poor showing of the country children were partly the result of their poor schooling, then the data obtained with the Primer scale should show a much greater per cent of the six, seven and eight year old country children above the medians for *their* age. As a matter of fact, only 22% of these younger country school children score above the median for their age, as determined from city children!

It would seem reasonable to conclude, then, that the differences found by both scales, between country and city children, were real differences in intelligence. The findings tally well with the frequent assertion of sociologists that the more intelligent individuals in the farming communities are constantly moving to the cities. It should be said in this connection that the country district studied is of a distinctly poor character, the land being hilly and unproductive, and many of the people being "poor whites" from the mountains of Kentucky.

The second group of data obtained with the Primer scale consists of results from 337 children six, seven and eight years old—all the children of these ages—in the schools of an Indiana city of about 12,000 inhabitants. The Primer scale was given these children in the first three grades as part of a total survey of the school system—the "Cross-Out" scale being given to children above the third grade. These last results (using again the data from children 10-14 years old—a total of 548 cases), were then grouped according to the occupation of the fathers. Children whose fathers were professional men (doctors, teachers, ministers, lawyers) were placed in one group; children whose fathers were executives (independent business men, foremen) constituted a second group. The third group was made up of the children of "artisans" (skilled workmen, machinists, railroad engineers); and the fourth group consisted of the children of unskilled laborers. The per cent, in each group, scoring above the median for their age, was then determined. These per cents ran as follows:

Occupation Group:	Professional	Executive	Artisan	Laborer
No. of cases	57	105	138	248
% above median for age	85	68	41	39

It was felt, however, that these findings might not be truly a measure of differences in native endowment, because of the possible influence of the environmental factors mentioned above. As a check on these factors, results obtained with the Primer scale from the children 6-8 years old were, therefore, grouped in the same way, and the groups compared by the same method. The per cents were as follows:

Occupation Group:	Professional	Executive	Artisan	Laborer
No. of cases	21	51	127	138
% above median for age	79	60	54	38

The results are again largely similar to the results obtained from the older children.[4]

It might be argued, however, that the two scales were testing, in different form, somewhat the same special abilities: particularly is this argument plausible since both scales used in this last study (of the correlation of occupation with intelligence) are "cross-out" scales, and are largely similar in general scheme, presentation and problem. To obtain some light upon this question the results with the Primer scale were, therefore, analyzed by test.

The first test of the scale consists of groups of dots, each group of dots making a pattern; there is, however, one dot in each group which is outside of or spoils the pattern—this dot the children are to cross out. The second test is made up of squares, each square containing two objects which are similar, in some important way, and one object which is different from the other two: this different thing the children are to cross out. The third test shows at the top of the page four forms—a triangle, a square, a cross, and a circle. Below are squares or "boxes" containing "blocks" which are to be fitted into the four places at the top of the page: but in each "box" there is one piece which will not fit in—and this extra piece is to be crossed out. The last test shows pictures in each of which there is something wrong—this wrong part is to be crossed out.

---

[4] It should be pointed out that the occupational groups are of different sizes, and that the smallness of the professional groups, in both instances, besides lowering the reliability, decreases the dispersion of the distribution and so exaggerates the superiority of this group as measured from the median of the entire group. For comparative purposes, however, as used here, the method serves well enough.

The second and last tests thus involve considerable information: these two tests might, then, be influenced by home environment. But it is hard to see how such environmental factors could operate to as great an extent in developing a child's ability to recognize dot patterns, or assist him in discriminating the geometrical forms of the third test. It should, therefore, be possible, by comparing results on the four tests, to form something of an estimate of the "reliability" of the measures obtained: in so far as the findings are constant from test to test of the scale used, it might be possible even to infer that similar findings would be obtained if other tests still were employed. The following table shows the per cent of children in each occupational group scoring above the median for their age, on each of the four tests:

Occupational Group:	Test 1	2	3	4
Professional	68	70	72	71
Executive	62	62	58	61
Artisan	50	61	51	54
Laborer	40	38	42	47

As will be seen, the per cents are distinctly constant from test to test. The differences are quite as great on the two tests which we would expect least influenced by home environment as they are on the two tests we would expect most sensitive to such influences.

Since then, the findings obtained from the two scales are consistent, and the findings are consistent from test to test, it seems reasonable to infer that the differences found between the occupational groups are probably true differences in a fundamental, underlying general intelligence or native endowment. And if we may assume that intelligence is hereditary we may argue back to the conclusion (regarding the parents of these children) that there is a positive correlation between occupational level and native ability—in general, at least, people find the level of work for which they are capable.

III. *Discussion.* The writer does not wish to press these conclusions. It is evident that all these tests are pencil and paper tests. The country children, young as well as older, are more shy with strangers than are city children. The children from well-to-do homes often have nursery games somewhat analogous to the tests.

But in so far as the results *are* valid, as thus interpreted, they have further interesting bearings upon certain general problems of mental measurement which should be pointed out. In the first place, these successive agreements between the scales used contribute to the validation of both types of examination. The scales for use with the older children stand out as, for the most part, untouched by environmental factors which might very well be expected to influence, in an illicit way, the findings. The substantial agreement of the Primer scale with the Standard tests included in the scales for the upper grades is evidence to show that the Primer scale is also measuring "general intelligence." More interesting still, however, is the way in which there emerges, from these various comparisons, the outline of a unitary "general ability." The data is obviously most inadequate as the basis for any inferences regarding such a large problem. But the writer cannot but feel that a large number of such simple and direct comparisons, between a variety of groups and using a variety of tests, are more needed at present in the study of general and special ability than more elaborate mathematical analyses of a relatively small amount of data.

Summary

The paper reports comparisons of (a) country and city children, and (b) children from different economic levels, by means of a group scale of intelligence applicable to the first three grades. It was found that

(1) 22% of the country children 6-8 years old score above the median for their age made by city children.

(2) Children of professional and business men rate distinctly above children of laboring men and mechanics.

(3) Similar results were found in surveys by means of scales applicable to the older children. It is, therefore, argued (a) that these differences previously found were differences in innate ability, not in schooling or home culture, and (b) that there was some general factor (presumably general mental endowment) independent of the particular tests used, with respect to which these groups differed.

VII. FIRST REVISION OF A GROUP SCALE DESIGNED FOR INVESTIGATING THE EMOTIONS, WITH TENTATIVE NORMS.

By S. L. PRESSEY AND O. R. CHAMBERS

1. *The Tests.* In the June number of this JOURNAL[1] certain "tests" were described intended for the investigation of emotional interests and distractibility. The tests were shortly after given to three small groups, one consisting of college students, another composed of girls from the State Industrial School, and a third made up of dementia praecox cases from two hospitals for the insane. The data were decidedly meager, but served nevertheless to indicate faults so marked as to make it seem wise to revise the tests before experimenting with them further. The tests as thus reconstructed may be described briefly as follows:

Test I. Affective Spread and Displacement: The test consists of 25 lists of words; each list contains five words, making a total of 125 words in all. All but 25 of the words name things more or less unpleasant. The subjects are told to read through the list and cross out every word that is unpleasant to them. There is no time limit, every subject being given time to finish. After the last subject has finished the last line the directions are that the group is to go through the list again, and draw a line around the one word in each line which is *most* unpleasant. In scoring, the total number of words marked unpleasant is first counted and used as a measure of affective spread or tendency to emotionalize. The number of lines in which the subject chooses as most unpleasant a word other than the word so chosen by the most of the average cases (that is, the modal word) is then counted and the sum used as a measure of emotional peculiarity or displacement.[2]

---

[1] S. L. Pressey and L. W. Pressey, "Cross-out" Tests, with Suggestions as to a Group Scale of the Emotions, *Journal of Applied Psychology*, Vol. III, 1919, pp. 138-150.

[2] Those who are familiar with Freudian terminology and theories will understand at once from the name of the test the general notion back of it, and the type of abnormal mental condition to which it is hoped the test will be sensitive. In fact, the five tests of the "scale" might, not altogether inaptly, be described as an attempt to investigate Freudianism experimentally.

The first five lines of the test run as follows:

1. disgust fear sex suspicion aunt
2. roar divorce dislike sidewalk wiggle
3. naked snicker wonder spit fight
4. failure home rotting snake hug
5. prize gutter thunder breast insult

Test II. Emotional Distractibility: The test consists of two parts: the first half is a paragraph of very commonplace and stupid reading matter, with 20 irrevelant words scattered in through the test; the subject is to read through the passage and cross out all irrelevant words. A rigid time limit of one minute and twenty seconds is set. At the end of that time the subjects are told to stop and go through the second paragraph in the same way. This second paragraph is sown in the same way with 20 irrelevant words. But it is a very gruesome description of a medieval execution. The score consists of the number of extra words missed in the first passage less the number of extra words missed in the second, the idea being that the emotional excitement of the second passage should cause the subject to overlook more irrelevant words here.[3]

The first five lines of each passage are given below:

This evening's the " Herald " says that the Milton property east of 3rd Street was sold this man morning to Smith and Cooper out of Chicago. It seems that is Smith has been, for some time, looking for a good poor piece of land in the business heart part of town upon rock which he might build another boat of his chain of 10c stores.

In the past time the most horrible and terrible forms of punishing crime were far common. Taunton tells many of an execution for treason under the most cruel and revolting in conditions. The man was hanged for three minutes, then, when his struggling wits began to decrease, was cut down, stripped, and his abdomen wide.

Test III. Moral Discrimination and Experience: The test is superficially somewhat similar to test I. It consists of 25 lists, each of five moral terms. The subjects are told to go through the lists and cross out in each list the thing that they consider worst. There is no time limit. The subjects are next told to go through the list again and draw a line around the wrong act or idea which they consider most common. The scores consist of the number of deviations from the most commonly chosen worst act and the most frequent sin. The idea has been that it might be illuminating to set over against each other moral and intellectual or experiential judgments in this way. As a matter of fact the judgments as to most common sin have proven most interesting.[4]

---

[3] The test was developed on the basis of some unpublisheed work done by one of the writers (Dr. Pressey) some years ago at the Psychopathic hospital with dementia praecox and psychopathic personality cases, but dates back ultimately to a card sorting test with pictures for distraction described by Boring (Boring, E. G. "Learning in Dementia Praecox" *Psychological Monographs*, Vol. 15, 1913, No. 63, pp. 101).

The first five lines of the test run as follows:

1. insulting, quarreling, mislead, hurting, carefulness.
2. borrowing, stealing, gambling, honesty, begging.
3. hate, rudeness, liking, dislike, fighting.
4. drunkenness, temperance, cursing, flirting, beating.
5. religious, crossness, smoking, stealing, swearing.

Test IV. Free Association: The test consists of a list of 25 words in capitals, each word in capitals being followed by a list of five other words in small letters. The subjects are told simply to go through the lists and draw a line through the one word in each list which is most closely connected in their minds with the word in big letters at the beginning of the list—they are to cross out the word which they most naturally think of in connection with the first word. There is no time limit. The score consists simply of the number of variations from the most common associate.[5]

---

[4] The test is an attempt to put in reasonably satisfactory and objective form an ethical discrimination test. It is therefore related to tests of this general nature described by Healy, Tests for Mental Classification *Psychological Monographs,* No. 2, Vol. 12, 1911 and Guy Fernald, The Defective Delinquent Class: Differentiating Tests, *American Journal of Insanity,* Vol. 68, No. 4, April, 1912.

[5] The test derives directly from the Kent-Rosanoff article (Kent, Grace Helen and Rosanoff, A. J., A study of Association in Insanity, *American Journal of Insanity,* Vol. 67, Nos. 1 and 2, 1910). The list of words in capitals is from the Kent-Rosanoff list of 100 words and most of the other words used are from their list of associates. As will be seen, the list begins with one association of very high frequency, according to their tables, in each line and goes down to absolute irrelevancy as near as the writers could make it.

It should also be mentioned that the writers have asked, after the associations have been marked, that the subjects go through the lists again and draw a line around each word that means "something to eat or drink, or something to wear, or a part of the body;" there are 25 such words in the lists. This part of the test was planned as merely a rough intelligence test for check on the general mental level of the groups which might be investigated. It is perhaps better omitted; the omission of this part may affect results on the last test, however.

The first five lines of the test run as follows:

1. BLOSSOM nice flower pour poison cheese
2. LAMP fear cheer match light dogs
3. BATH nakedness hen water danger paper
4. KING dog tyrant fish queen grade
5. SLEEP midnight rest beautiful worry baseball

Test V. Emotional Memory: The test consists of a list of one hundred words, fifty of which have occurred in the previous tests, and fifty of which have not. Of these fifty, twenty-five have been chosen as emotional and twenty-five as unemotional. The subjects are told to go through the list and cross out all the words which they think have occurred in the previous tests. Two scores are obtained: (a) excess of emotional over unemotional words correctly remembered, and (b) excess of emotional over unemotional words which were marked as remembered but which did not occur in the previous test.[6]

The first five lines of the test run as follows:

fear finger paper story nude rose east brutal slashed
liquid business hacked crave Smith hate screamed piano
ground dollars author belly ripped rock yards pears flirting
railroad vomit cow horrible seduce trust mind gloomy water
lover funeral tall rape shrieked parts pencil ghastly

2. *Nature of the Revision.* The general nature of the revision may be very briefly indicated. The four tests first experimented with have already been described in the previous article referred to above (note 1). These tests were given to some thirty college students, and about the same number of girls at the state reform school and dementia praecox cases at two state hospitals.[7] Study of results from these three groups led to the following general conclusions. In the first place, tests for use with such widely different groups (and particularly group tests for use with the insane)

---

[6] The notion of the test derives from the experience of one of the writers (Dr. Pressey) as subject for an experiment by Tolman; among other subjects Tolman investigated the influence of affective toning upon memory (Tolman, E. C., and Isabelle Johnson, *Am. J. Psych.*, Vol. 29, 1918, pp. 187-195, but the subject, is of course an old one with a very considerable literature.

[7] Acknowledgments are due to Dr. Kenosha Sessions, superintendent of the Indiana Girls' School, to Dr. Max Bahr, of the Central Indiana Hospital for the Insane, and to Miss Hazel Hansford, psychologist and field worker at the Southeastern Indiana Hospital for the Insane, for their kindness and help in securing the data.

cannot be time-limit tests. In the second place, every effort must be made to use words which will be known to everyone; limitations of vocabulary must be carefully considered, particularly in working with delinquent groups. In the third place, some method is desirable by which response in chance fashion to a test may be found out; otherwise, in scoring papers from psychotics, it is often impossible to tell erratic chance reactions from erratic thinking which is nevertheless on the problem.

As a result of these considerations only one time limit test was kept in the revised form presented above. Comparison of girls' school and college results served to indicate at least the most gross differences in vocabulary. And a check on chance reaction was sought by the use of "jokers." That is, in each line of the "what is worst" test there is one virtue, and in each line of the "unpleasant" test there is one word which is either positively pleasant or at least not marked as unpleasant by any one in either the college or the girl's school group.[8] In scoring the tests these jokers are first glanced over, and any records showing an appreciable number of responses on the jokers thrown out. In an effort to obtain the maximum amount of information from each test the first, third, and fourth tests are made to yield double scores; the device appears to work very well (in fact, the writers seriously considered making one test yield three scores) and would seem of some general usefulness.

The changes made in the individual tests cannot be gone into in detail. It may be said shortly that the data obtained from the three groups mentioned above were analyzed very closely, elements which appeared differential seized upon and more like them added, and the sensitivity of the tests increased by proper grouping of items. The materials were also arranged more systematically, and so as to permit of more ready analysis. One test was dropped altogether, one test of the "Cross-out" scale[9] radically made over and included in this series, and the free association test added.

---

[8] In these two tests the jokers are put into the test according to a set scheme, in order to facilitate scoring; the joker is fifth in the first line, fourth in the second line, third in the third line, and so on, the series beginning over again in the sixth line. This scheme is not readily hit upon by a subject. And it is also (a more important point) not the sort of thing a subject might fall in with unwittingly as the result of an automatism; a subject might react uniformly to the first word in every line, or the middle word, he might take progressively the next word, but he would hardly be likely to work backward in this way. These considerations are particularly important in the first test, where not only the jokers, but all the classifications, run in this way.

The general nature of the changes made may be illustrated by the revision of the first test. The first results made it evident that there are certain words, such as murder, which practically everybody,—psychotic, delinquent, or college student,—consider unpleasant; other words such as smile everyone considers pleasant. Such words are evidently of no value (except as "jokers"); they are not differential. Therefore in the final form no word of the first test was kept unless it was considered, by more than 20% and less than 80% of both the college group *and* the girls' school group, to be unpleasant; words of about the same percentage of unpleasantness were also put together in the same line in order to make the test sensitive to small difference of opinion or affective attitude. But it was felt that this was not enough. It was to be expected that the delinquents and psychotics would differ from normal folk not merely by more random choices; the atypical cases would show instead peculiar but consistent trends. A guess was therefore made as to what trends would be interesting to investigate, and words of five types chosen, words which would be unpleasant because of their relation to disgust, fear, sex, or self-feeling, and the jokers. One of each type of word was put in each line, after the scheme used with the jokers; the key to the arrangement is given by the first line of the test, as shown above.

3. *The Tentative "Norms."* The tests, as thus revised, were then given to a total of 101 college students, 49 men and 52 women. From these results the following tentative norms were worked out:

Test I. (a) per cent of each sex, and of the entire group, considering each word unpleasant; (b) number of deviations, in choice of the most unpleasant word, from the modal choice.

Test II. differences between number of extra words correctly crossed out in unemotional and emotional passages unemotional less emotional).

Test III. (a) per cent, for each sex and for the entire group, considering each wrong act worst, and most common; (b) number of deviations, in choice of the worst and the most unpleasant, from the modal choice.

---

[9] See the previous article mentioned in note 1 above.

Test IV. (a) per cent, for each sex and for the entire group, choosing each word as an associate; (b) number of deviations, in choice of the associate, from the modal choice.[10]

Test V. (a) excess (or the reverse) of emotional over unemotional words correctly remembered; (b) excess (or the reverse) of emotional over unemotional words marked, but not actually occurring in the previous tests.

4. *Purpose of the Tests.* The writers realize, very decidedly, the crudeness of the tests, and the inadequacy of the data accumulated so far. But the data already obtained will serve at least for a rough first orientation in dealing with any further material which may later be accumulated. Results from a group of factory hands or other relatively unskilled laborers, from a group of colored adults, a group of delinquents, and a pathological group (preferably neurotics or early dementia praecox) are particularly desired. The writers are not so situated that such data are readily obtainable; it is with the hope that others, who are already working with such special groups, may be interested to thus experiment that the present paper is being published. The test forms, and tabulations to date, will be gladly furnished to any who may be interested to do such work.

And as an indication of what can be done in the development of such data the writers wish to present very briefly, in closing, the results of an analysis, for sex differences, of the responses on the first test. The subjects were told to cross out the words which were unpleasant to them. It was found that 55% of the women marked more words as unpleasant than did the median man—that is, the difference was negligible. However, the per cent, for each sex, marking each word as unpleasant was next found, the twelve most differential words located, and the number of *these* words marked as unpleasant by each man and each woman counted. It was found that 94% of the women marked more of *these* words unpleasant than did the median man!

We do not, of course need mental tests for the distinguishing of the sexes; and it may be said that the writers have tried, so far as possible, to avoid items on which sex differences might appear so that separate sex norms would not be necessary.[11] But this bit of analysis will show something of the possibilities of development contained in the tests; the writers see no reason for doubting that differential groups within each test, which shall prove little less effective, may be found in working on neurotics and psychopathic delinquents. At least there is sufficient evidence to make the experiment seem worth while.

---

[10] The number of correct responses, in making the "things to eat or drink, things to wear, and parts of the body" were also tabulated; but the measure would seem of little valuo.

A CORRECTION

In a minor study entitled "The Efficiency of the Group Point Scale in Prognosticating Success and Failure in Junior High School," in the December number of this Journal, a prognosis chart or percentage correlation table for prognostic purposes was presented. By some mischance this chart was confused with a similar chart for correlation of score with Teachers' Estimates. The correlation coefficient given in the article was correct. But the chart should read as follows:

		Division into fifths—test score—				
		1	2	3	4	5
Division into fifths according to school marks	V	0	0	30	20	50
	IV	15	20	20	25	20
	III	15	25	15	30	15
	II	35	25	20	10	10
	I	35	30	15	15	5

In each array one is lowest and five highest.

S. L. Pressey.

---

[11] For the most part this has been accomplished. Thus on the fourth test the two sexes differ in only two instances, in their selection of the worst sin. But nevertheless only 22% of the women differ as much as the median man, from the modal selection—a result due perhaps to wider experience on the part of the men. It is also interesting that half again more sex words were selected by the women than by the men as most unpleasant, and that the men exceeded the women in selection of the "fear" words.

It should be added in this connection that the tests are by no means wholly a masculine production; each writer frequently consulted his wife, and the majority of the items of the first form were selected by Mrs. Pressey,—whose suggestions and help throughout the entire course of the work have been of the greatest value.

BOOK REVIEWS

C. E. SEASHORE. *The Psychology of Musical Talent.* Silver, Burdett and Company, Boston, 1919, p. xvi+288.

In this volume Professor Seashore has assembled his tests on musical talent, which he presents with typical results in simple and untechnical language. Though the volume is addressed to students of applied psychology, the author has evidently had also in mind the music supervisor with little psychological training, for one misses a searching analysis of the data collected, and the treatment can hardly be called systematic though a very definite scheme of classification underlies the work.

Musical talent is considered under five heads: musical sensitivity, musical action, musical memory and imagination, musical intellect and musical feeling. As a basis of analysis the author accepts four auditory attributes—pitch, intensity, duration and extensity. Tests for the first three of these "senses" are described and their elemental importance magnified. "Pitch is to the musician what color is to the artist—his medium of expression." Accordingly a test of pitch-discrimination is taken to be a basic test of musical ability. The experiments seem to indicate that the physiological limit can be attained after a brief practice period and that it does not change with age or further training, nor vary with respect to sex. If pitch discrimination alone is poor we can predict a corresponding inferiority in all its derived factors—though good pitch discrimination does not insure excellence in these other factors. As a result of testing large numbers for pitch-discrimination, Seashore has reached the following practical conclusions. A person who discriminates 3 vibrations or better at the level of 435 d.v. may become a musician. One whose discrimination falls between 3 and 8 d.v. should have a plain musical education; between 9 and 17 d.v. one should have it only if a special inclination for some kind of music is shown; one whose discrimination requires 18 vibrations or more should have nothing to do with music.

A normal distribution curve indicates the percentage of persons giving varying percentages of right judgments, the largest number, something over 30 per cent, giving 85 per cent of right judgments in these tests. Tests are also recorded for the range of pitch, the results showing increased sensitivity from the level of 64 d. v. to 128 d. v. with an at first gradual and then more rapid decrease from 256 d.v. upwards.

With regard to the "sense" of intensity, interferences in hearing are briefly discussed together with the phenomena of tonal gaps and tonal islands. The use of the audiometer is described and also the pitch-range audiometer for intensities at different pitches. The standard of acuity for a very good ear allows the faintest sound of the instrument to be heard at each pitch-level from 200 to 3200 d. v., but there is no indication that these intensities are either subjectively or objectively equal. If they were, the results would be at variance with the normal curve, established by Max Wien.[1] Little difference is found between the acuity of adults and children. No improvement was apparent after training and blind persons were found to possess no better hearing than persons with normal vision.

In studying the sense of time the author assumes a motor theory of response. Discrimination as fine as 1/100 of a second is extraordinary, while a record as poor as 1/2 of a second is equally rare. Though children do less well than adults in these tests, improvement with practice is attributed to a growing knowledge of the time-process. The very slight correlation of precision in hearing the durations of tones or of short intervals with lessons in music leads to the conclusion that either musical training does not improve a capacity so elemental as the time sense, or that those with a good time sense are no more apt than others to be selected for a musical education. Pitch and time are found to be fairly independent variables; the need of a good time sense in music being regarded as contingent.

The sense of rhythm is reduced to five fundamental capacities; the senses of time and of intensity, auditory and motor imagery and a motor impulse for rhythm; its measurement is correspondingly reduced to these elements. Timbre is described as a complex of pitches. The criteria for judgments of consonance are listed as *blending*, *smoothness* and *purity*. The instruction for the consonance test is to "give the decison on blending alone if the degree of blending (in the two-clang comparison) is perceptibly different; if not make the decision on smoothness—and, if there is no difference in either smoothness or blending base the decision on purity." A normal order resulting from experiments with the piano is recorded and two comparison tests, a longer and a shorter, are described. The results of these tests show them to be independent of the age of observer and, while improvement with training is evident, it is not so noticeable as one might suppose. The test may, therefore, be used to advantage before the child knows anything of music.

In the discussion of auditory space, the ability of the ears singly and together in determining the direction of a source of sound is described and explained in terms of difference in intensity. The sense of extensity "varies exactly parallel with pitch; there is an inseparable duality." "For the purpose of rating talent it is therefore unnecessary to concern ourselves with the isolation of the sense"; a conclusion which must seem of doubtful validity to those who have followed the recent studies of hearing; particularly Rich's investigations of the volume threshold which demonstrate that its limen is quite different from that of pitch.

Seashore differentiates volume and extensity, preferring to use the former term as a combined effect involving several factors such as extensity, intensity, timbre and reduplication of sound.

On the subject of motor control, the basic motor capacities are outlined with respect to time as motility, timed action, response to a simple signal, action upon choice and serial action; with respect to movement; precision, discrimination, strength and endurance. A variety of tests are described employing the chronograph and chronoscope, the simple tapping test being adopted as an index to motility.

Under the head of musical action Seashore describes the use of his tonoscope, and some of the chief results secured with it in measuring voluntary control of pitch.

In considering musical imagery and imagination a somewhat exaggerated importance is attached to the auditory image. The results of a questionnaire indicated that "as a rule musicians who rate themselves low offer some excuse, explaining that they suffer from defective capacity in this particular; that they have neglected to develop it; that they have been engaged in some sort of musical business which does not make it necessary, etc." Whatever importance may attach to imagery it can hardly be doubted that the specific imagery of a special sense is less significant in indicating a particular talent of appreciation or execution than was once supposed to be the case; and musicians who deplore their lack of auditory imagery may be unduly impressed when their attention has been called to some apparent deficiency in their ability to hear inwardly.

---

[1] Pflüger's Arch. f. d. ges. Physiol., 1903, 97, pp. 1 ff.

As for musical imagination Seashore's treatment is suggestive rather than precisely analytical. "Imagination proper" he writes, "is not a specific mental process, as is sensation or perception, but is rather a designation for certain group functions of images, associations, thoughts, feelings, and efforts in countless permutations." A statement equally vague deals with the language of music: "It is to the credit of language if it convey one specific idea and that only; it is to the credit of music if it lead to a richer self-expression transcending the bounds of defined concepts and literal form." In the opinion of the reviewer, music transcends "the bounds of defined concepts" no more than does verbal language, but the definition of the musical concept as distinct from certain of the elements of auditory experience is lacking in this work.

Of the remaining chapters the one on musical memory describes tests for memory-span and retention, and includes a discussion of learning curves and absolute pitch. With respect to musical intellect a sound statement is made that "musical thought is a specialization in dealing with the problems which arise in music. Although the form and content of the thought are different, it requires the same kind of logical grasp as in mathematics or philosophy." Yet no effort is made to clarify the peculiar form and content of music, as might perhaps have been done by reference to the work of Stumpf, Lipps, M. F. Meyer and W. V. Bingham. Indeed Seashore's conclusions rest almost exclusively upon results obtained in his own laboratory. The chapter on musical feeling is equally vague.

The final chapter deals with pedagogical hints and the advocacy of a consulting supervisor of music who may be competent to test the musical talents of children in the schools and to judge them in accordance with their capacities and interests.

For all the technical ingenuity with which Seashore has attacked the important problem of musical diagnosis one feels that somehow the chief feature of musical talent has escaped his "dragnet;" this being the ability to conceive and to think in terms of the musical interval. The neglect of the interval casts a doubt upon his "basic" test of pitch. In discriminating pitch within the range of 3 vibrations at a level of 435 d.v. we are not dealing with a musical interval but with a mere difference of "height." The difference between a semitone above 435 d.v. in just intonation (464 d.v.) and the same note in tempered intonation (460 d.v.) is greater than the limit which Seashore's test makes the basis of a musical diagnosis. The reviewer believes that the attitudes wherein one judges *height* and *interval* are quite different and that they rest upon different "senses." In view of Rich's work on the volume threshold it seems highly probable that the independent attribute or "sense of extensity" which Seashore neglects as being "exactly parallel with pitch" is nevertheless the elemental foundation upon which our judgments of interval and the logic of music are based. In the report on pitch discrimination at different levels of pitch, we find a curve which indicates that in terms of fractional parts of a tone keenness in the sense of pitch remains approximately constant from 256 d. v. to 2048 d. v. This curve therefore parallels Rich's limens in that a certain fraction of the vibration appears to correlate with a just noticeable difference of "pitch," as does a similar fraction with a just noticeable difference of volume. It would thus appear possible that Seashore's pitch test involves both pitch-height and pitch-interval. Some observers judging pitch-height may discriminate differences of 3 d. v. and better at the level of 435; while others influenced by the attitude of interval-difference may tend to require differences exceeding the limen of volume, which at this level would be about 9 d. v.

How far this neglect of a direct test for the sense of interval may impair the results of Seashore's diagnosis, it is not easy to say. A corrective is of course to be found in several of the accessory tests which measure the sense of interval indirectly, such as those for consonance and those employing the tonoscope, in singing intervals and in voice control; but the emphasis which the author places upon the test of pitch can not seem just to those who define pitch as height, and who must therefore deny that "pitch is to the musician what color is to the artist—his medium of expression."

Cornell University. R. M. OGDEN.

J. W. BRIDGES. *An Outline of Abnormal Psychology*. R. G. Adams & Co., Columbus, Ohio, 1919, p. 127.

This book is intended to be useful to "those medical students and students of social service who desire a general survey of this field but who have insufficient time for a regular supervised course or for extensive reading of the very much scattered literature." It is also "to serve as a guide for students of abnormal psychology in the absence of a comprehensive text-book."

The book is found, however, to be more than a mere guidebook. For example, in the second part, on mental diseases, a concise and carefully arranged table of symptoms is given under the discussion of each type of insanity. The name, "An Outline of Abnormal Psychology" may be taken literally: the book is written actually in outline form. The first part deals with abnormal phenomena in general. Under definitions and classification, Wernicke's classification only is given. This, however, is not held to in the author's discussion of the subject, which begins with sensation, goes on through consciousness and attention, memory, association, judgment, orientation feeling and temperament, to instinct and emotion, innate action, and acquired action, much in the manner of the ordinary text-book of normal psychology. The section ends with chapters on intelligence, personality, and sleep, dreams, and hypnosis. Each chapter contains a full list of all possible abnormalities in the different manifestations of mind, together with their technical nomenclature. Under intelligence, the most generally used methods of measuring intelligence are mentioned. The second part is on mental disease. Feeble-mindedness, usually given at least a chapter in a work on abnormal psychology, is not treated here, except as one kind (cretinism) finds a place under thyroigenous psychoses. All the other varying forms, such as mongolianism, and microcephalis which certainly ought to be described in a book on abnormal psychology, (at least if that term is taken in its broadest sense to mean not-normal psychology) are not even mentioned but are summarily dismissed when the author distinguishes between amentia or feeble-mindedness, and dementia, explaining that "The former is an innate defect, while the latter is the result of a mental deterioration." No general classifications of the psychoses are given, except the author's own, as implied by his chapter headings, which run as follows: Dementia Precox, (Paraphrenia, Paranoia), Manic-Depressive Insanity, Dementia Paralytica, The Alcoholic Psychoses, Morphine, Cocaine, and Other Drugs, The Presenile and Senile Psychoses, and The Symptomatic Psychoses. But in the third part, on Borderline Diseases, Freud's, Sidis's, and Kraepelin's classifications of the neuroses are presented.

Throughout the work impartial consideration is given to many and widely differing theories of the various phases of abnormal psychology. Copious references are to be found after each chapter. One wishes that a comprehensive index might also have been given a place.

This outline will probably be of greater assistance to those already fairly familiar with the subject, but who need help in clarifying and systematizing their knowledge, than to those who, by the reading of this book alone, would gain their whole information. The work is at once too technical and too abbreviated to be very meaningful to the average laymen. MARJORY BATES.

Clark University.

CARTER ALEXANDER. *School Statistics and Publicity.* Silver, Burdett & Co., New York, 1919, p. xix+332.

This little book is an outgrowth of the author's teaching at the Peabody College for Teachers. It is produced with the avowed aim of aiding active superintendents of schools to adequately place before the public such statistics as show school needs and school achievements. It is also intended for use as a textbook by those who are "engaged in training future superintendents." Its clear, simple style and the abundance of illustrative material make it very serviceable for either purpose, and as a textbook its value is enhanced by the suggestions to instructors and by the exercises which follow each chapter.

The book falls rather readily into three main divisions. The first of these, consisting of the first three chapters is largely introductory in value. The author points out in the first chapter the various types of errors and deficiencies commonly noted when good but unthinking men are tempted to use figures. The second chapter very practically tells how to collect data, and the third is an argument for knowledge of the technique of statistics on the part of school administrators.

There can be no doubt in the mind of the reader after reading these three chapters that the author thoroughly believes in statistical, or at least graphic presentation of school facts to the public. At the same time he takes a very moderate stand on the question of the amount of statistical knowledge that is practically necessary.

The second division, consisting of chapters four to eight inclusive, affords a very clear elementary presentation of statistical method. The fifth, sixth and seventh chapters are the most difficult in the book and the author admits that they will require little more careful study than the others, but despite this admission it would be hard to find a clearer or simpler exposition of the matters treated. Chapters six and seven, treating of measures of deviation and of relationship, are the most technical, and it is possible that in a majority of cases the busy superintendent could omit them without serious loss so far as their value in connection with his local public is concerned. If he is to take his work very seriously, however, they are very needful in order that he may be able to interpret his data completely.

The remaining four chapters, forming the third division, present devices for making clear and striking presentations of data. These devices are well chosen, not only by reason of their effectiveness but also because of their ingenious use of materials and tools ready at hand even in a poorly equipped office. This practicality is further evidenced by the author's statement that all but five of the seventy-nine cuts in the book were made by high school boys,—a demonstration of the possibilities of student assistance to the superintendent.

An annotated bibliography is given, consistent in type with the rest of the book, its chief features being its freedom from cumbrousness and its choice, in the main, of readily accessible material.

Clark University. GEO. ALLEN COE.

WILLIAM H. DOOLEY. *Principles and Methods of Industrial Education For Use in Teacher Training Classes.* With an Introduction by Charles A. Prosser) Houghton Mifflin Company, Boston, 1919, p. 257.

Probably no educational scheme or problem is more in the public eye than that of vocational or industrial education. The paucity of skilled help in nearly all industries is such that the problem of present supply is a most perplexing and serious one. Dr. Prosser sums up the situation in his Introduction to this book: "Never in our history has there been such a keen realization of the dependence of production upon skill, and the part that wise methods of training have in cultivating skill. In our ways and means for meeting these increasing demands we are at once fortunate and unfortunate; fortunate in adequate financial support for sound instructor training plans; unfortunate in a shortage of people to organize and direct them, and doubly unfortunate in a lack of organized practical material for use in instructor (teacher) training classes."

This book, as a whole, goes far to fill the place which Dr. Prosser notes above, and is an admirable compilation of contemporary writers and bulletins, and is drawn without stint from any source that seemed to furnish the material for the purpose in mind.

It has been difficult for the vocational teacher-training personnel to get away from the academic point of view, and the sporadic efforts of the past and even of the present, lean too far away from practical shop sense to meet the need of the industries for practical, not thoretical, men. The author has seemingly been fortunate to sense this most important error and has arranged his book in orderly fashion, and clothed it in understandable language that easily makes interesting many of the dry details of pedagogy which are essential for the well-trained trade teacher.

The book deals with the following topics: the Value of Industrial Education, the Educational Needs of Trades and Industries, How men have been trained for Trades and Industries in the Past, Different Types of Industrial Schools, Organization of Industrial Schools and of Industrial Classes, an Industrial Survey, Principles of Psychology underlying Learning, General Methods of Teaching, General Methods for Teaching in Industrial Education, Methods of Teaching Shop Work, of Interpretation of Blue Prints and Shop Sketching, of Shop

Science, of Shop or Industrial Mathematics and of English and Manual Training compared with Industrial Education.

A valuable feature of the book is the list of "Questions for Discussion" that is given at the end of each chapter. This is followed by a "List of References for Further Reading." There is also an Appendix of forty-seven pages giving typical outlines and suggestive courses of study in various kinds of vocational schools.

Taken as a whole, the work should prove valuable to the trade instructor and especially to those engaged in teacher-training for vocational schools.

Clark University. HARRY E. MILLIKEN.

NOTES

The Carnegie Corporation of New York has announced its purpose to give $5,000,000 for the use of the National Academy of Sciences and the National Research Council. It is understood that a portion of the money will be used to erect in Washington a home of suitable architectural dignity for the two beneficiary organizations. The remainder will be placed in the hands of the Academy, which enjoys a federal charter, to be used as a permanent endowment for the National Research Council. This impressive gift is a fitting supplement to Mr. Carnegie's great contributions to science and industry.

The Council is a democratic organization based upon some forty of the great scientific and engineering societies of the country, which elect delegates to its constitutent Divisions. It is not supported or controlled by the government, differing in this respect from other similar organizations established since the beginning of the war in England, Italy, Japan, Canada, and Australia. It intends, if possible to achieve in a democracy and by democratic methods the great scientific results which the Germans achieved by autocratic methods in an autocracy while avoiding the obnoxious features of the autocratic regime.

The Council was organized in 1916 as a measure of national preparedness and its efforts during the war were mostly confined to assisting the government in the solution of pressing war-time problems involving scientific investigation. Reorganized since the war on a peace-time footing, it is now attempting to stimulate and promote scientific research in agriculture, medicine, and industry, and in every field of pure science. The war afforded a convincing demonstration of the dependence of modern nations upon scientific achievement, and nothing is more certain than that the United States will ultimately fall behind in its competition with the other great peoples of the world unless there be persistent and energetic effort expended to foster scientific discovery.

The Commissioner of Institutions and Agencies of New Jersey, Burdette G. Lewis, has recently extended the scope of the State Psychiatric Clinic to include psychological work in the correctional institutions of New Jersey. The work is at present organized as a section of the Psychiatric Clinic of the New Jersey State Hospital for the Insane, Dr. Henry A. Cotton, Director. Mr. Edgar A. Doll was appointed as the Department Psychologist in July, 1919, following a period of preliminary investigation concerning the application of psychological methods in correctional institutions. In August, 1919, Mr. W. J. Ellis was appointed assistant psychologist. The psychological staff also includes 2 volunteer assistants, Mr. Warren S. Prince and Miss Anna Gillingham. The psychological work in the correctional institutions of New Jersey is directly related to the classification and marking system recently installed as a basis for parole under the Division of Education and Parole of the Department of Institutions and Agencies, Calvin Derrick, Director. The functions of the psychological work are to make mental diagnoses of the inmates of correctional institutions and on the basis of psychological tests and measurements to recommend concerning the educational, vocational, disciplinary and parole treatment of delinquents and prisoners. The Army group test Alpha, supplemented by individual examinations, has been in use since February, 1919, and surveys of the four principal institutions have been made. Clinical psychological methods to supplement the group tests have been installed in the New Jersey State Prison, at the State Home for Boys at Jamesburg, and at the State Home for Girls at Trenton. Psychological analyses of industrial and vocational activities of the institutions are in process at the State Prison and at the State Home for Boys. It is expected that if the work continues to develop the Staff of the Psychological Section will be expanded as needs demand.

The present program of psychological examining at the New Jersey State Prison includes an attempt to introduce scientific management in the problems of industrial and vocational assignments of prisoners. The purpose of this program is (1) to improve the possibilities of utilizing to the best advantage the reformatory influences which the Prison provides, (2) to improve the morale or mental attitude of the prisoners by having them assigned to work which is suited to their capabilities and best interests, (3) to improve the efficiency of instruction in the vocational shops, and (4) to increase the effectiveness of the routine work in the Prison shops.

To accomplish these ends the psychological section of the Psychiatric Clinic has instituted at the request of the Commissioner and the Director of the Division of Education and Parole, a psychological and industrial analysis of the Prison activities both vocational and industrial. This analysis is intended to form the basis of intelligent assignment of men to shops or tasks. The regular work of the Psychologist calls for individual analysis of each prisoner with respect to his general intelligence, mental responsibility, vocational aptitudes, and industrial qualifications. It is the hope of the Department to effect the coordination of this man-analysis and job-analysis in such a way that, knowing the capabilities of the man and the requirements of the job the welfare of the prisoners and their efficiency in their tasks may be materially advanced.

Job-analysis. The Assistant Psychologist aided by a graduate student in psychology has conducted psychological and vocational analyses of numerous 'jobs' in the Prison. The results of this investigation cannot yet be presented in full but the results for two jobs may be cited briefly. The Assistant Psychologist finds that the print shop, for example, is a highly specialized industry, calling for particular degrees of skill which are ordinarily attained only after long practice. The print shop also presents an industry whose various operations are highly differentiated with respect to the gradation of processes in the industry. The Psychologist, therefore, has found very clear dependence of success upon general intelligence and general education in the succession of tasks in the print shop. It is, therefore,, comparatively easy to predict a man's degree of success in learning this industry simply on the basis of his general intelligence and general education. In addition it is possible in the print shop to measure the degree of specific aptitude and skill which men must possess for immediate success in the work of the shop. It is easily possible to apply mental tests such as alphabet sorting and visual discrimination, and tests of special forms of fatigue which would be applicable to the assignment of men to work in the print shop.

Contrasted with the print shop we may cite briefly the results obtained in the shoe shop. In this shop the succession of tasks is not very sharply differentiated on the basis of general intelligence. The general level of intelligence in the shoe shop ranges from a minimum of mental age 11 years to a maximum mental age of about 14 years. The overlapping of intelligence from operation to operation is very great. Moreover, the educational requirements are practically negligible. Therefore, in the shoe shop it is difficult to assign a man to a particular task on the basis of his general intelligence or education. In the print shop, by contrast, the range of intelligence was from his minimum mental age of 11 years to a maximum mental age of 18 years. The succession of operations within that shop were found to be very clearly operated on the basis of general intelligence. In the shoe shop as in the print shop it is possible to devise tests of specific aptitude for the several operations demanded. A cutter, for example, must be able to cut up a piece of leather with a minimum of waste. A psychological test of form perception analogous to the jig-saw puzzle might be used for testing this ability. As yet the psychologist has not proceeded far enough with his work to devise the tests needed for such work. He is, however, at work on this problem.

Man-analysis. Each prisoner is examined by the psychologist almost immediately after his admission to the Prison. The purpose of this examination is to determine general intelligence, character or personality make up, actual degree of education and potential educability, actual degree of industrial skill and potential vocational aptitudes. The psychological and educational tests are relatively well developed. It is therefore easy enough to measure a man's general intelligence, education and degree of educability with a high degree of accuracy. Trade and vocational tests, however, are not yet sufficiently developed for immediate use at the Prison. Those tests for measuring industrial capability and vocational aptitudes which are now available in psychology are not specifically suited to the needs of the work at the Prison. It is therefore necessary to devise special tests of this character specifically suited to the present Prison conditions. It is planned to obtain the industrial history of each man and then to quiz him carefully regarding his present degree of skill in the major trades which he claims to have followed. It is also planned to examine him with vocational tests to determine his industrial aptitudes. He will also be questioned concerning his own ambitions and the industrial possibilities in the environment to which he expects to go when released.

Coordination of Job-analysis and Man-analysis. Specification cards will be designed for all the detailed jobs in the Prison and similar cards will be designed to summarize all the information obtainable regarding each man. These cards will call for the physical, mental, educational and industrial requirements of the job and the similar abilities of the man. It will then be possible, knowing the requirements of the job and the capabilities of the man, to assign a man to a job on a scientific basis. It is, of course, unlikely that this relatively ideal program can be achieved within the course of at least another year, but some actual progress is already being made. The assigning officer at the Prison is already using the results of the psychological examining in such a way as to enable him to place men more successfully than heretofore. The general conduct of the work involved in carrying out this program is also definitely bringing to consciousness among the Prison officials both the needs and the possibilities for more efficient classification and placement.

THE
JOURNAL OF APPLIED PSYCHOLOGY
VOL. IV JUNE-SEPTEMBER, 1920 Nos. 2 and 3

TABLES TO FACILITATE THE COMPUTATION OF COEFFICIENTS OF CORRELATION BY THE RANK DIFFERENCE METHOD*

These tables have been prepared to facilitate the computation of the coefficient of correlation by the rank difference method.

Correlation coefficients based on ranks are necessarily used when the variables to be correlated are expressed in ranks. Certain writers have failed to see that in such cases a coefficient based on ranks in the variables is just as significant as one based on measurements. When the interest is in relative position the rank difference correlation coefficient should not be converted into r; ρ gives the actual relationship in terms of relative position.

Computation of correlation coefficients by the rank difference method is frequently useful even when original measurements must first be transformed into ranks. If the relation is not desired in terms of relative position, the rank correlation coefficient may be converted into r by means of the table on page 125, provided the original distribution can be considered approximately normal.

There is an impression that the rank formula is an approximation. This is not true. The formula

$$\rho = 1 - \frac{6\Sigma D^2}{n(n^2 - 1)} \text{ is derived directly from } r = \frac{\Sigma(xy)}{N\sigma_x\,\sigma_y} \text{ and}$$

corresponds exactly with the product moment correlation coefficient.

These tables will be found useful in eliminating some of the drudgery of computation in scientific laboratories and in classroom work in statistics, psychology, and the other sciences which make use of correlation statistics. Industry will also find in these tables a short cut to the determination of many important relationships. This is especially true of work carried on in employment departments.

THE SCOTT COMPANY LABORATORY.

Philadelphia, Pa.
March 1, 1920.

---

*Reprints of this article with the Tables may be obtained from the *Journal of Applied Psychology*. Orders should be addressed to Miss Florence Chandler, Clark University, Worcester, Mass.

INSTRUCTIONS FOR CALCULATING RANK DIFFERENCE COEFFICIENTS WITH THE USE OF TABLES*

1. Arrange each of the two series of measurements to be correlated in rank order. Every pair of measurements is thus represented by a pair of ranks — the relative positions assigned those measurements in the two series.

2. Obtain the difference between the two ranks assigned each case.

3. Square each of these differences. The squares are read from the Table, Page 117. Note that the decimals may be omitted from the squares without appreciable error.

4. Add these squares of the differences. This sum (ΣD^2) is to be found in the body of the Tables, Pages 118 to 125.

5. Refer in the Tables to the proper column for the number of cases used (N). Thus, if 21 cases were ranked, use the second column on Page 119. Run down this column until you find the entry nearest the obtained ΣD^2.

6. Read the entry in this same line at either side of the page (first or last column). This entry is ρ, the desired coefficient of correlation.

Rank in Series I	Rank in Series II	Difference Squared
1	7	36
2	9	49
3	2	1
4	12	64
5	5	0
6	3	9
7	8	1
8	1	49
9	15	36
10	11	1
11	13	4
12	4	64
13	16	9
14	19.5	30.25
15	6	81
16	18	4
17	14	9
18	10	64
19	21	4
20	19.5	.25
21	17	16
	$\Sigma D^2 =$	531.5

Example: 21 cases are arranged in rank order for two variables; the sum of the squares of the differences in the corresponding ranks is found to be 531 (five-tenths is disregarded).

Refer to the second column on Page 119, run down to the entry which most nearly equals 531, that is 524 and read beside it the coefficient of correlation, .66.

Note: (a) If the obtained ΣD^2 is nearly midway between two entries, use the ρ falling between the ρ's corresponding to the two entries. (If the ΣD^2 in the example were 540, ρ would be .65.) Correlation coefficients accurate to .01 can be read almost instantly. Closer interpolation is useless.

(b) For ΣD^2 greater than those corresponding to $\rho = 0$ (i. e. in case of negative correlation) subtract the obtained ΣD^2 from the entry in that column corresponding to $\rho = -1.00$. Find the ρ corresponding to this number obtained by subtraction. With a negative sign prefixed, this is the desired ρ.

---

*For a complete discussion of the rank difference method of computing correlation coefficients, see H. O. Rugg, "Statistical Methods Applied to Education," P. 283 ff., Houghton Mifflin Co., New York; E. L. Thorndike, "Introduction to the Theory of Mental and Social Measurements," P. 24 ff, P. 157 ff, Teachers College, Columbia University, New York; Wm. Stern, "Psychological Methods of Testing Intelligence," P. 109 ff, Warwick & York, Baltimore, Md.

SQUARES OF DIFFERENCES

From 0 to 80 by Halves

Difference	Square of difference	Difference	Square of difference	Difference	Square of difference	Difference	Square of difference
0	0	20	400.00	40	1600.00	60	3600.00
.5	.25	20.5	420.25	40.5	1640.25	60.5	3660.25
1	1.00	21	441.00	41	1681.00	61	3721.00
1.5	2.25	21.5	462.25	41.5	1722.25	61.5	3782.25
2	4.00	22	484.00	42	1764.00	62	3844.00
2.5	6.25	22.5	506.25	42.5	1806.25	62.5	3906.25
3	9.00	23	529.00	43	1849.00	63	3969.00
3.5	12.25	23.5	552.25	43.5	1892.25	63.5	4032.25
4	16.00	24	576.00	44	1936.00	64	4096.00
4.5	20.25	24.5	600.25	44.5	1980.25	64.5	4160.25
5	25.00	25	625.00	45	2025.00	65	4225.00
5.5	30.25	25.5	650.25	45.5	2070.25	65.5	4290.25
6	36.00	26	676.00	46	2116.00	66	4356.00
6.5	42.25	26.5	702.25	46.5	2162.25	66.5	4422.25
7	49.00	27	729.00	47	2209.00	67	4489.00
7.5	56.25	27.5	756.25	47.5	2256.25	67.5	4556.25
8	64.00	28	784.00	48	2304.00	68	4624.00
8.5	72.25	28.5	812.25	48.5	2352.25	68.5	4692.25
9	81.00	29	841.00	49	2401.00	69	4761.00
9.5	90.25	29.5	870.25	49.5	2450.25	69.5	4830.25
10	100.00	30	900.00	50	2500.00	70	4900.00
10.5	110.25	30.5	930.25	50.5	2550.25	70.5	4970.25
11	121.00	31	961.00	51	2601.00	71	5041.00
11.5	132.25	31.5	992.25	51.5	2652.25	71.5	5112.25
12	144.00	32	1024.00	52	2704.00	72	5184.00
12.5	156.25	32.5	1056.25	52.5	2756.25	72.5	5256.25
13	169.00	33	1089.00	53	2809.00	73	5329.00
13.5	182.25	33.5	1122.25	53.5	2862.25	73.5	5402.25
14	196.00	34	1156.00	54	2916.00	74	5476.00
14.5	210.25	34.5	1190.25	54.5	2970.25	74.5	5550.25
15	225.00	35	1225.00	55	3025.00	75	5625.00
15.5	240.25	35.5	1260.25	55.5	3080.25	75.5	5700.25
16	256.00	36	1296.00	56	3136.00	76	5776.00
16.5	272.25	36.5	1332.25	56.5	3192.25	76.5	5852.25
17	289.00	37	1369.00	57	3249.00	77	5929.00
17.5	306.25	37.5	1406.25	57.5	3306.25	77.5	6006.25
18	324.00	38	1444.00	58	3364.00	78	6084.00
18.5	342.25	38.5	1482.25	58.5	3422.25	78.5	6162.25
19	361.00	39	1521.00	59	3481.00	79	6241.00
19.5	380.25	39.5	1560.25	59.5	3540.25	79.5	6320.25

Decimals may be entirely disregarded without appreciably affecting the resulting correlation coefficients. The error is never greater than .02, and is usually much less.

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	17	18	19	20	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	00	00	00	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	19	23	27	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	39	46	53	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	58	68	80	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	78	91	106	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	97	114	133	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	116	137	160	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	136	160	186	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	155	182	213	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	174	205	239	.8
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	194	228	266	.8
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	213	251	293	.7
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	233	274	319	.7
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	252	296	346	.7
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	271	319	372	.7
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	291	342	399	.70
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	310	365	426	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	330	388	452	.66
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	349	410	479	.64
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	368	433	505	.62
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	388	456	532	.60
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	407	479	559	.58
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	426	502	585	.56
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	446	524	612	.54
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	465	547	638	.52
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	485	570	665	.50
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	504	593	692	.48
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	523	616	718	.46
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	543	638	745	.44
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	562	661	771	.42
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	581	684	798	.40
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	601	707	825	.38
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	620	730	851	.36
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	640	752	878	.34
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	659	775	904	.32
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	571	678	798	931	.30
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	698	821	958	.28
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	717	844	984	.26
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	736	866	1011	.24
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	756	889	1037	.22
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	775	912	1064	.20
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	795	935	1091	.18
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	814	958	1117	.16
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	702	833	980	1144	.14
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	718	853	1003	1170	.12
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	734	872	1026	1197	.10
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	751	891	1049	1224	.08
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	767	911	1072	1250	.06
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	783	930	1094	1277	.04
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	800	950	1117	1303	.02
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	816	969	1140	1330	.00
[illegible]	440	572	728	910	1120	1360	1632	1938	2280	2660	—1.00

NUMBER OF CASES RANKED

ρ	21	22	23	24	25	26	27	28	29	30	ρ
1.00	00	00	00	00	00	00	00	00	00	00	1.00
.98	31	35	40	46	52	59	66	73	81	90	.98
.96	62	71	81	92	104	117	131	146	162	180	.66
.94	92	106	121	138	156	176	197	219	244	270	.94
.92	123	142	162	184	208	234	262	292	325	360	.92
.90	154	177	202	230	260	293	328	365	406	450	.90
.88	185	213	243	276	312	351	393	438	487	539	.88
.86	216	248	283	322	364	410	459	512	568	629	.86
.84	246	283	324	368	416	468	524	585	650	719	.84
.82	277	319	364	414	468	527	590	658	731	809	.82
.80	308	354	405	460	520	585	655	731	812	899	.80
.78	339	390	445	506	572	644	721	804	893	989	.78
.76	370	425	486	552	624	702	786	877	974	1079	.76
.74	400	460	526	598	676	761	852	950	1056	1169	.74
.72	431	496	567	644	728	819	917	1023	1137	1259	.72
.70	462	531	607	690	780	878	983	1096	1218	1349	.70
.68	493	567	648	736	832	936	1048	1169	1299	1438	.68
.66	524	602	688	782	884	995	1114	1242	1380	1528	.66
.64	554	638	729	828	936	1053	1179	1315	1462	1618	.64
.62	585	673	769	874	988	1112	1245	1389	1543	1708	.62
.60	616	708	810	920	1040	1170	1310	1462	1624	1798	.60
.58	647	744	850	966	1092	1229	1376	1535	1705	1888	.58
.56	678	779	891	1012	1144	1287	1441	1608	1786	1978	.56
.54	708	815	931	1058	1196	1346	1507	1681	1868	2068	.54
.52	739	850	972	1104	1248	1404	1572	1754	1949	2158	.52
.50	770	886	1012	1150	1300	1463	1638	1827	2030	2248	.50
.48	801	921	1052	1196	1352	1521	1704	1900	2111	2337	.48
.46	832	956	1093	1242	1404	1580	1769	1973	2192	2427	.46
.44	862	992	1133	1288	1456	1638	1835	2046	2274	2517	.44
.42	893	1027	1174	1334	1508	1697	1900	2119	2355	2607	.42
.40	924	1063	1214	1380	1560	1755	1966	2192	2436	2697	.40
.38	955	1098	1255	1426	1612	1814	2031	2265	2517	2787	.38
.36	986	1133	1295	1472	1664	1872	2097	2338	2598	2877	.36
.34	1016	1169	1336	1518	1716	1931	2162	2412	2680	2967	.34
.32	1047	1204	1376	1564	1768	1989	2228	2485	2761	3057	.32
.30	1078	1240	1417	1610	1820	2048	2293	2558	2842	3147	.30
.28	1109	1275	1457	1656	1872	2106	2355	2631	2923	3236	.28
.26	1140	1311	1498	1702	1924	2165	2424	2704	3004	3326	.26
.24	1170	1346	1538	1748	1976	2223	2490	2777	3086	3416	.24
.22	1201	1381	1579	1794	2028	2282	2555	2850	3167	3506	.22
.20	1232	1417	1619	1840	2080	2340	2621	2923	3248	3596	.20
.18	1263	1452	1660	1886	2132	2399	2686	2996	3329	3686	.18
.16	1294	1488	1700	1932	2184	2457	2752	3069	3410	3776	.16
.14	1324	1523	1741	1978	2236	2516	2817	3142	3492	3866	.14
.12	1355	1558	1781	2024	2288	2574	2883	3216	3573	3956	.12
.10	1386	1594	1822	2070	2340	2633	2948	3289	3654	4046	.10
.08	1417	1629	1862	2116	2392	2691	3014	3362	3735	4135	.08
.06	1448	1665	1903	2162	2444	2750	3079	3435	3816	4225	.06
.04	1478	1700	1943	2208	2496	2808	3145	3508	3898	4315	.04
.02	1509	1736	1984	2254	2548	2867	3210	3581	3979	4405	.02
.00	1540	1771	2024	2300	2600	2925	3276	3654	4060	4495	.00

NUMBER OF CASES RANKED

ρ	11	12	13	14	15	16	17	18	19	20	ρ
1.00	00	00	00	00	00	00	00	00	00	00	1.00
.98	4	6	7	9	11	14	16	19	23	27	.98
.96	9	11	15	18	22	27	33	39	46	53	.96
.94	13	17	22	27	34	41	49	58	68	80	.94
.92	18	23	29	36	45	54	65	78	91	106	.92
.90	22	29	36	46	56	68	82	97	114	133	.90
.88	26	34	44	55	67	82	98	116	137	160	.88
.86	31	40	51	64	78	95	114	136	160	186	.86
.84	35	46	58	73	90	109	131	155	182	213	.84
.82	40	51	66	82	101	122	147	174	205	239	.82
.80	44	57	73	91	112	136	163	194	228	266	.80
.78	48	63	80	100	123	150	180	213	251	293	.78
.76	53	69	87	109	134	163	196	233	274	319	.76
.74	57	74	95	118	146	177	212	252	296	346	.74
.72	62	80	102	127	157	190	228	271	319	372	.72
.70	66	86	109	137	168	204	245	291	342	399	.70
.68	70	92	116	146	179	218	261	310	365	426	.68
.66	75	97	124	155	190	231	277	329	388	452	.66
.64	79	103	131	164	202	245	294	349	410	479	.64
.62	84	109	138	173	213	258	310	368	433	505	.62
.60	88	114	146	182	224	272	326	388	456	532	.60
.58	92	120	153	191	235	286	343	407	479	559	.58
.56	97	126	160	200	246	299	359	426	502	585	.56
.54	101	132	167	209	258	313	375	446	524	612	.54
.52	106	137	175	218	269	326	392	465	547	638	.52
.50	110	143	182	228	280	340	408	485	570	665	.50
.48	114	149	189	237	291	354	424	504	593	692	.48
.46	119	154	197	246	302	367	441	523	616	718	.46
.44	123	160	204	255	314	381	457	543	638	745	.44
.42	128	166	211	264	325	394	473	562	661	771	.42
.40	132	172	218	273	336	408	490	581	684	798	.40
.38	136	177	226	282	347	422	506	601	707	825	.38
.36	141	183	233	291	358	435	522	620	730	851	.36
.34	145	189	240	300	370	449	539	640	752	878	.34
.32	150	194	248	309	381	462	555	659	775	904	.32
.30	154	200	255	319	392	476	571	678	798	931	.30
.28	158	206	262	328	403	490	588	698	821	958	.28
.26	163	212	269	337	414	503	604	717	844	984	.26
.24	167	217	277	346	426	517	620	736	866	1011	.24
.22	172	223	284	355	437	530	636	756	889	1037	.22
.20	176	229	291	364	448	544	653	775	912	1064	.20
.18	180	235	298	373	459	558	669	795	935	1091	.18
.16	185	240	306	382	470	571	685	814	958	1117	.16
.14	189	246	313	391	482	585	702	833	980	1144	.14
.12	194	252	320	400	493	598	718	853	1003	1170	.12
.10	198	257	328	410	504	612	734	872	1026	1197	.10
.08	202	263	335	419	515	626	751	891	1049	1224	.08
.06	207	269	342	428	526	639	767	911	1072	1250	.06
.04	211	275	349	437	538	653	783	930	1094	1277	.04
.02	216	280	357	446	549	666	800	950	1117	1303	.02
.00	220	286	364	455	560	680	816	969	1140	1330	.00
1.00	440	572	728	910	1120	1360	1632	1938	2280	2660	—1.00

NUMBER OF CASES RANKED

ρ	21	22	23	24	25	26	27	28	29	30	ρ
1.00	00	00	00	00	00	00	00	00	00	00	1.00
.98	31	35	40	46	52	59	66	73	81	90	.98
.96	62	71	81	92	104	117	131	146	162	180	.66
.94	92	106	121	138	156	176	197	219	244	270	.94
.92	123	142	162	184	208	234	262	292	325	360	.92
.90	154	177	202	230	260	293	328	365	406	450	.90
.88	185	213	243	276	312	351	393	438	487	539	.88
.86	*216*	*248*	*283*	*322*	*364*	*410*	*459*	*512*	*568*	*629*	*.86*
.84	246	283	324	368	416	468	524	585	650	719	.84
.82	277	319	364	414	468	527	590	658	731	809	.82
.80	308	354	405	460	520	585	655	731	812	899	.80
.78	339	390	445	506	572	644	721	804	893	989	.78
.76	370	425	486	552	624	702	786	877	974	1079	.76
.74	400	460	526	598	676	761	852	950	1056	1169	.74
.72	431	496	567	644	728	819	917	1023	1137	1259	.72
.70	462	531	607	690	780	878	983	1096	1218	1349	.70
.68	493	567	648	736	832	936	1048	1169	1299	1438	.68
.66	524	602	688	782	884	995	1114	1242	1380	1528	.66
.64	554	638	729	828	936	1053	1179	1315	1462	1618	.64
.62	585	673	769	874	988	1112	1245	1389	1543	1708	.62
.60	616	708	810	920	1040	1170	1310	1462	1624	1798	.60
.58	647	744	850	966	1092	1229	1376	1535	1705	1888	.58
.56	678	779	891	1012	1144	1287	1441	1608	1786	1978	.56
.54	708	815	931	1058	1196	1346	1507	1681	1868	2068	.54
.52	739	850	972	1104	1248	1404	1572	1754	1949	2158	.62
.50	770	886	1012	1150	1300	1463	1638	1827	2030	2248	.50
.48	801	921	1052	1196	1352	1521	1704	1900	2111	2337	.48
.46	832	956	1093	1242	1404	1580	1769	1973	2192	2427	.46
.44	862	992	1133	1288	1456	1638	1835	2046	2274	2517	.44
.42	893	1027	1174	1334	1508	1697	1900	2119	2355	2607	.42
.40	924	1063	1214	1380	1560	1755	1966	2192	2436	2697	.40
.38	955	1098	1255	1426	1612	1814	2031	2265	2517	2787	.38
.36	986	1133	1295	1472	1664	1872	2097	2338	2598	2877	.36
.34	1016	1169	1336	1518	1716	1931	2162	2412	2680	2967	.34
.32	1047	1204	1376	1564	1768	1989	2228	2485	2761	3057	.32
.30	1078	1240	1417	1610	1820	2048	2293	2558	2842	3147	.30
.28	1109	1275	1457	1656	1872	2106	2355	2631	2923	3236	.28
.26	1140	1311	1498	1702	1924	2165	2424	2704	3004	3326	.26
.24	1170	1346	1538	1748	1976	2223	2490	2777	3086	3416	.24
.22	1201	1381	1579	1794	2028	2282	2555	2850	3167	3506	.22
.20	1232	1417	1619	1840	2080	2340	2621	2923	3248	3596	.20
.18	1263	1452	1660	1886	2132	2399	2686	2996	3329	3686	.18
.16	1294	1488	1700	1932	2184	2457	2752	3069	3410	3776	.16
.14	1324	1523	1741	1978	2236	2516	2817	3142	3492	3866	.14
.12	1355	1558	1781	2024	2288	2574	2883	3216	3573	3956	.12
.10	1386	1594	1822	2070	2340	2633	2948	3289	3654	4046	.10
.08	1417	1629	1862	2116	2392	2691	3014	3362	3735	4135	.08
.06	1448	1665	1903	2162	2444	2750	3079	3435	3816	4225	.06
.04	1478	1700	1943	2208	2496	2808	3145	3508	3898	4315	.04
.02	1509	1736	1984	2254	2548	2867	3210	3581	3979	4405	.02
.00	1540	1771	2024	2300	2600	2925	3276	3654	4060	4495	.00

NUMBER OF CASES RANKED

ρ	31	32	33	34	35	36	37	38	39	40	ρ
1.00	00	00	00	00	00	00	00	00	00	00	1.00
.98	99	109	120	131	143	155	169	183	198	213	.98
.96	198	218	239	262	286	311	337	366	395	426	.96
.94	298	327	359	393	428	466	506	548	593	640	.94
.92	397	436	479	524	571	622	675	731	790	853	.92
.90	496	546	598	655	714	777	844	914	988	1066	.90
.88	595	655	718	785	857	932	1012	1097	1186	1279	.88
.86	694	764	838	916	1000	1088	1181	1279	1383	1492	.86
.84	794	873	957	1047	1142	1243	1350	1462	1581	1706	.84
.82	893	982	1077	1178	1285	1399	1518	1645	1778	1919	.82
.80	992	1091	1197	1309	1428	1554	1687	1828	1976	2132	.80
.78	1091	1200	1316	1440	1571	1709	1856	2011	2174	2345	.78
.76	1190	1309	1436	1571	1714	1865	2025	2193	2371	2558	.76
.74	1290	1419	1556	1702	1856	2020	2193	2376	2569	2772	.74
.72	1389	1528	1676	1833	1999	2176	2362	2559	2766	2985	.72
.70	1488	1637	1795	1964	2142	2331	2531	2742	2964	3198	.70
.68	1587	1746	1915	2094	2285	2486	2700	2924	3162	3411	.68
.66	1686	1855	2035	2225	2428	2642	2868	3107	3359	3624	.66
.64	1786	1964	2154	2356	2570	2797	3037	3290	3557	3838	.64
.62	1885	2073	2274	2487	2713	2953	3206	3473	3754	4051	.62
.60	1984	2182	2394	2618	2856	3108	3374	3657	3952	4264	.60
.58	2083	2292	2513	2749	2999	3263	3543	3838	4150	4477	.58
.56	2182	2401	2633	2880	3142	3419	3712	4021	4347	4690	.56
.54	2282	2510	2753	3011	3284	3574	3881	4204	4545	4904	.54
.52	2381	2619	2872	3142	3427	3730	4049	4387	4742	5117	.52
.50	2480	2728	2992	3273	3570	3885	4218	4570	4940	5330	.50
.48	2579	2837	3112	3403	3713	4040	4387	4752	5138	5543	.48
.46	2678	2946	3231	3534	3856	4196	4555	4935	5335	5756	.46
.44	2778	3055	3351	3665	3998	4351	4724	5118	5593	5970	.44
.42	2877	3164	3471	3796	4141	4507	4893	5301	5730	6183	.42
.40	2976	3274	3590	3927	4284	4662	5062	5483	5928	6396	.40
.38	3075	3383	3710	4058	4427	4817	5230	5666	6126	6609	.38
.36	3174	3492	3830	4189	4570	4973	5399	5849	6323	6822	.36
.34	3274	3601	3949	4320	4712	5128	5568	6032	6521	7036	.34
.32	3373	3710	4069	4451	4855	5284	5736	6215	6718	7249	.32
.30	3472	3819	4189	4582	4998	5439	5905	6397	6916	7462	.30
.28	3571	3928	4308	4712	5141	5594	6074	6580	7114	7675	.28
.26	3670	4037	4428	4843	5284	5750	6243	6763	7311	7888	.26
.24	3770	4147	4548	4974	5426	5905	6411	6946	7509	8102	.24
.22	3869	4256	4668	5105	5569	6061	6580	7128	7706	8315	.22
.20	3968	4365	4787	5236	5712	6216	6749	7311	7904	8528	.20
.18	4067	4474	4907	5367	5855	6371	6918	7494	8102	8741	.18
.16	4166	4583	5027	5498	5998	6527	7086	7677	8299	8954	.16
.14	4266	4692	5146	5629	6140	6682	7255	7860	8497	9168	.14
.12	4365	4801	5266	5760	6283	6838	7424	8042	8694	9381	.12
.10	4464	4910	5386	5891	6426	6993	7592	8225	8892	9594	.10
.08	4563	5020	5505	6021	6569	7148	7761	8408	9090	9807	.08
.06	4662	5129	5625	6152	6712	7304	7930	8591	9287	10020	.06
.04	4762	5238	5745	6283	6854	7459	8099	8773	9485	10234	.04
.02	4861	5347	5864	6414	6997	7615	8267	8956	9682	10447	.02
.00	4960	5456	5984	6545	7140	7770	8436	9139	9880	10660	.00
—1.00	9920	10912	11968	13090	14280	15540	16872	18278	19760	21320	—1.00

NUMBER OF CASES RANKED

ρ	41	42	43	44	45	46	47	48	49	50	ρ
1.00	00	00	00	00	00	00	00	00	00	00	1.00
.98	230	247	265	284	304	324	346	368	392	417	.98
.96	459	494	530	568	607	649	692	737	784	833	.96
.94	689	740	795	851	911	973	1038	1105	1176	1250	.94
.92	918	987	1060	1135	1214	1297	1384	1474	1568	1666	.92
.90	1148	1234	1324	1419	1518	1622	1730	1842	1960	2083	.90
.88	1378	1481	1589	1703	1822	1946	2076	2211	2352	2499	.88
.86	1607	1728	1854	1987	2125	2270	2421	2579	2744	2916	.86
.84	1837	1975	2119	2270	2429	2594	2767	2948	3136	3332	.84
.82	2066	2221	2384	2554	2732	2919	3113	3316	3528	3749	.82
.80	2296	2468	2649	2838	3036	3243	3459	3685	3920	4165	.80
.78	2526	2715	2914	3122	3340	3567	3805	4053	4312	4582	.78
.76	2755	2962	3179	3406	3643	3892	4151	4422	4704	4998	.76
.74	2985	3209	3443	3689	3947	4216	4497	4790	5096	5415	.74
.72	3214	3455	3708	3973	4250	4540	4843	5159	5488	5831	.72
.70	3444	3702	3973	4257	4554	4865	5189	5527	5880	6248	.70
.68	3674	3949	4238	4541	4858	5189	5535	5896	6272	6664	.68
.66	3903	4196	4503	4825	5161	5513	5881	6264	6664	7081	.66
.64	4133	4443	4768	5108	5465	5837	6227	6633	7056	7497	.64
.62	4362	4690	5033	5392	5768	6162	6572	7001	7448	7914	.62
.60	4592	4936	5298	5676	6072	6486	6918	7370	7840	8330	.60
.58	4822	5183	5562	5960	6376	6810	7264	7738	8232	8747	.58
.56	5051	5430	5827	6244	6679	7135	7610	8107	8624	9163	.56
.54	5281	5677	6092	6527	6983	7459	7956	8475	9016	9580	.54
.52	5510	5924	6357	6811	7286	7783	8302	8844	9408	9996	.52
.50	5740	6171	6622	7095	7590	8108	8648	9212	9800	10413	.50
.48	5970	6417	6887	7379	7894	8432	8994	9580	10192	10829	.48
.46	6199	6664	7152	7663	8197	8756	9340	9949	10584	11246	.46
.44	6429	6911	7417	7946	8501	9080	9686	10317	10976	11662	.44
.42	6658	7158	7682	8230	8804	9405	10032	10685	11368	12079	.42
.40	6888	7405	7946	8514	9108	9729	10378	11054	11760	12495	.40
.38	7118	7651	8211	8798	9412	10053	10724	11423	12152	12912	.38
.36	7347	7898	8476	9082	9715	10378	11069	11791	12544	13328	.36
.34	7577	8145	8741	9365	10019	10702	11415	12160	12936	13745	.34
.32	7806	8392	9006	9649	10322	11026	11761	12528	13328	14161	.32
.30	8036	8639	9271	9933	10626	11351	12107	12897	13720	14578	.30
.28	8266	8886	9536	10217	10930	11675	12453	13265	14112	14994	.28
.26	8495	9132	9801	10501	11233	11999	12799	13634	14504	15411	.26
.24	8725	9379	10065	10784	11537	12323	13145	14002	14896	15827	.24
.22	8954	9626	10330	11068	11840	12648	13491	14371	15288	16244	.22
.20	9184	9873	10595	11352	12144	12972	13837	14739	15680	16660	.20
.18	9414	10120	10860	11636	12448	13296	14183	15108	16072	17077	.18
.16	9643	10366	11125	11920	12751	13621	14529	15476	16464	17493	.16
.14	9873	10613	11390	12203	13055	13945	14875	15845	16856	17910	.14
.12	10102	10860	11655	12487	13358	14269	15220	16213	17248	18326	.12
.10	10332	11107	11920	12771	13662	14594	15566	16582	17640	18743	.10
.08	10562	11354	12184	13055	13966	14918	15912	16950	18032	19159	.08
.06	10791	11601	12449	13339	14269	15242	16258	17319	18424	19576	.06
.04	11021	11847	12714	13622	14573	15566	16604	17687	18816	19992	.04
.02	11250	12094	12979	13906	14876	15891	16950	18056	19208	20409	.02
.00	11480	12341	13244	14190	15180	16215	17296	18424	19600	20825	.00

NUMBER OF CASES RANKED

ρ	51	52	53	54	55	56	57	58	59	60	ρ
1.00	00	00	00	00	00	00	00	00	00	00	1.00
.98	442	469	496	525	554	585	617	650	684	720	.98
.96	884	937	992	1049	1109	1170	1234	1300	1369	1440	.96
.94	1326	1406	1488	1574	1663	1756	1851	1951	2053	2159	.94
.92	1768	1874	1984	2099	2218	2341	2468	2601	2738	2879	.92
.90	2210	2343	2480	2624	2772	2926	3086	3251	3422	3599	.90
.88	2652	2811	2976	3148	3326	3511	3703	3901	4106	4319	.88
.86	3094	3280	3473	3673	3881	4096	4320	4551	4791	5039	.86
.84	3536	3748	3969	4198	4435	4682	4937	5201	5475	5758	.84
.82	3978	4217	4465	4722	4990	5267	6554	5852	6160	6478	.82
.80	4420	4685	4961	5247	5544	5852	6171	6502	6844	7198	.80
.78	4862	5154	5457	5772	6098	6437	6788	7152	7528	7918	.78
.76	5304	5622	5953	6296	6653	7022	7405	7802	8213	8638	.76
.74	5746	6091	6449	6821	7207	7608	8023	8452	8897	9357	.74
.72	6188	6559	6945	7346	7762	8193	8640	9103	9582	10077	.72
.70	6630	7028	7441	7871	8316	8778	9257	9753	10266	10797	.70
.68	7072	7496	7937	8395	8870	9363	9874	10403	10950	11517	.68
.66	7514	7965	8433	8920	9425	9948	10491	11053	11635	12237	.66
.64	7956	8433	8929	9445	9979	10534	11108	11703	12319	12956	.64
.62	8398	8902	9426	9969	10534	11119	11725	12353	13004	13676	.62
.60	8840	9370	9922	10494	11088	11704	12342	13004	13688	14396	.60
.58	9282	9839	10418	11019	11642	12289	12960	13654	14372	15116	.58
.56	9724	10307	10914	11543	12197	12874	13577	14304	15057	15836	.56
.54	10166	10776	11410	12068	12751	13460	14190	14954	15741	16555	.54
.52	10608	11244	11906	12593	13306	14045	14811	15604	16426	17275	.52
.50	11050	11713	12402	13118	13860	14630	15428	16255	17110	17995	.50
.48	11492	12182	12898	13642	14414	15215	16045	16905	17794	18715	.48
.46	11934	12650	13394	14167	14969	15800	16662	17555	18479	19435	.46
.44	12376	13119	13890	14692	15523	16386	17279	18205	19163	20154	.44
.42	12818	13587	14386	15216	16078	16971	17896	18855	19848	20874	.42
.40	13260	14056	14882	15741	16632	17556	18514	19505	20532	21594	.40
.38	13702	14524	15378	16266	17186	18141	19131	20156	21216	22314	.38
.36	14144	14993	15875	16790	17741	18726	19748	20806	21901	23034	.36
.34	14586	15461	16371	17315	18295	19312	20365	21456	22585	23753	.34
.32	15028	15930	16867	17840	18850	19897	20982	22106	23270	24473	.32
.30	15470	16398	17363	18365	19404	20482	21599	22756	23954	25193	.30
.28	15912	16867	17859	18889	19958	21067	22216	23406	24638	25913	.28
.26	16354	17335	18355	19414	20513	21652	22833	24057	25323	26633	.26
.24	16796	17804	18851	19939	21067	22238	23451	24707	26007	27352	.24
.22	17238	18272	19347	20463	21622	22823	24068	25357	26692	28072	.22
.20	17680	18741	19843	20988	22176	23408	24685	26007	27376	28792	.20
.18	18122	19209	20339	21513	22730	23993	25302	26657	28060	29512	.18
.16	18564	19678	20835	22037	23285	24578	25919	27308	28745	30232	.16
.14	19006	20146	21331	22562	23839	25164	26536	27958	29429	30951	.14
.12	19448	20615	21828	23087	24394	25749	27153	28608	30114	31671	.12
.10	19890	21083	22324	23612	24948	26334	27770	29258	30798	32391	.10
.08	20332	21552	22820	24136	25502	26919	28388	29908	31482	33111	.08
.06	20774	22020	23316	24661	26057	27504	29005	30558	32167	33831	.06
.04	21216	22489	23812	25186	26611	28090	29622	31209	32851	34550	.04
.02	21658	22957	24308	25710	27166	28675	30239	31859	33536	35270	.02
.00	22100	23426	24804	26235	27720	29260	30856	32509	34220	35990	.00
-1.00	44200	46852	49608	52470	55440	58520	61712	65018	68440	71980	—1.00

NUMBER OF CASES RANKED

ρ	61	62	63	64	65	66	67	68	69	70	ρ
1.00	00	00	00	00	00	00	00	00	00	00	1.0
.98	756	794	833	874	915	958	1002	1048	1095	1143	.9
.96	1513	1588	1667	1747	1830	1916	2005	2096	2190	2286	.9
.94	2269	2383	2500	2621	2746	2874	3007	3144	3284	3429	.9
.92	3026	3177	3333	3494	3661	3832	4009	4192	4379	4572	.9
.90	3782	3971	4166	4368	4576	4791	5012	5239	5474	5716	.9
.88	4538	4765	5000	5242	5491	5749	6014	6287	6569	6859	.8
.86	5295	5560	5833	6115	6406	6707	7016	7335	7664	8002	.8
.84	6051	6354	6666	6989	7322	7665	8019	8383	8758	9145	.8
.82	6807	7148	7500	7862	8237	8623	9021	9431	9853	10288	.8
.80	7564	7942	8333	8736	9152	9581	10023	10479	10948	11431	.8
.78	8320	8736	9166	9610	10067	10539	11026	11527	12043	12574	.7
.76	9077	9531	9999	10483	10982	11497	12028	12575	13138	13717	.7
.74	9833	10325	10833	11357	11898	12455	13030	13622	14232	14860	.7
.72	10590	11119	11666	12230	12813	13413	14032	14670	15327	16003	.7
.70	11346	11913	12499	13104	13728	14372	15035	15718	16422	17147	.7
.68	12102	12708	13332	13978	14643	15330	16037	16766	17517	18290	.6
.66	12859	13502	14166	14851	15558	16288	17039	17844	18612	19433	.6
.64	13615	14296	14999	15725	16474	17246	18042	18862	19706	20576	.6
.62	14372	15090	15832	16598	17389	18204	19044	19910	20801	21719	.6
.60	15128	15884	16666	17472	18304	19162	20046	20958	21896	22862	.6
.58	15884	16679	17499	18346	19219	20120	21049	22005	22991	24005	.5
.56	16641	17473	18332	19219	20134	21078	22051	23053	24086	25148	.5
.54	17397	18267	19165	20093	21050	22036	23053	24101	25180	26291	.5
.52	18154	19061	19999	20966	21965	22994	24056	25149	26275	27434	.5
.50	18910	19856	20832	21840	22880	23953	25058	26197	27370	28578	.5
.48	19666	20650	21665	22714	23795	24911	26060	27245	28465	29721	.4
.46	20423	21444	22499	23587	24710	25869	27063	28293	29560	30864	.4
.44	21179	22238	23332	24461	25626	26827	28065	29341	30654	32007	.4
.42	21936	23032	24165	25334	26541	27785	29067	30389	31749	33150	.4
.40	22692	23827	24998	26208	27456	28743	30070	31436	32844	34293	.4
.38	23448	24621	25832	27082	28371	29701	31072	32484	33939	35436	.3
.36	24205	25415	26665	27955	29286	30659	32074	33532	35034	36579	.3
.34	24961	26209	27498	28829	30202	31617	33077	34580	36128	37722	.3
.32	25718	27003	28332	29702	31117	32575	34079	35628	37223	38865	.3
.30	26474	27798	29165	30576	32032	33534	35081	36676	38318	40009	.3
.28	27230	28592	29998	31450	32947	34492	36084	37724	39413	41152	.2
.26	27987	29386	30831	32323	33862	35450	37086	38772	40508	42295	.2
.24	28743	30180	31665	33197	34778	36408	38088	39819	41602	43438	.2
.22	29500	30975	32498	34070	35693	37366	39090	40867	42697	44581	.2
.20	30256	31769	33331	34944	36608	38324	40093	41915	43792	45724	.2
.18	31012	32563	34164	35818	37523	39282	41095	42963	44887	46867	.1
.16	31769	33357	34998	36691	38438	40240	42097	44011	45982	48010	.1
.14	32525	34151	35831	37565	39354	41198	43100	54059	47076	49153	.1
.12	33282	34946	36664	38438	40269	42156	44102	46107	48171	50296	.1
.10	34038	35740	37498	39312	41184	43115	45104	47155	49266	51440	.1
.08	34794	36534	38331	40186	42099	44073	46107	48202	50361	52583	.0
.06	35551	37328	39164	41059	43014	45031	47109	49250	51456	53726	.0
.04	36307	38123	39997	41933	43930	45989	48111	50298	52550	54869	.0
.02	37064	38917	40831	42806	44845	46947	49114	51346	53645	56012	.0
.00	37820	39711	41664	43680	45760	47905	50116	52394	54740	57155	.0
−1.00	75640	79422	83328	87360	91520	95810	100232	104788	109480	114910	1.0

NUMBER OF CASES RANKED

ρ	ρ	71	72	73	74	75	76	77	78	79	80	ρ
1.00	1.00	00	00	00	00	00	00	00	00	00	00	1.00
.98	.98	1193	1214	1296	1350	1406	1463	1521	1581	1643	1706	.98
.96	.96	2386	2488	2593	2701	2812	2926	3043	3163	3286	3413	.96
.94	.94	3578	3732	3889	4052	4218	4389	4565	4745	4930	5119	.94
.92	.92	4771	4976	5186	5402	5624	5852	6086	6326	6573	6826	.92
.90	.90	5964	6220	6482	6753	7030	7315	7608	7908	8216	8532	.90
.88	.88	7157	7464	7779	8103	8436	8778	9129	9489	9859	10238	.88
.86	.86	8350	8707	9075	9454	9842	10241	10651	11071	11502	11945	.86
.84	.84	9542	9951	10372	10804	11248	11704	12172	12653	13146	13651	.84
.82	.82	10735	11195	11668	12155	12654	13167	13694	14234	14789	15358	.82
.80	.80	11928	12439	12965	13505	14060	14630	15215	15816	16432	17061	.80
.78	.78	13121	13683	14261	14856	15466	16093	16737	17397	18075	18770	.78
.76	.76	14314	14927	15558	16206	16872	17556	18258	18979	19718	20477	.76
.74	.74	15506	16171	16854	17557	18278	19019	19780	20561	21362	22183	.74
.72	.72	16699	17415	18151	18907	19684	20482	21301	22142	23005	23890	.72
.70	.70	17892	18659	19447	20258	21090	21945	22823	23724	24648	25596	.70
.68	.68	19085	19903	20744	21608	22496	23408	24344	25305	26291	27302	.68
.66	.66	20278	21147	22040	22959	23902	24871	25866	26887	27934	29009	.66
.64	.64	21470	22391	23337	24309	25308	26434	27387	28468	29578	30715	.64
.62	.62	22663	23634	24633	25660	26714	27797	28909	30050	31221	32422	.62
.60	.60	23856	24878	25930	27010	28120	29260	30430	31632	32864	34128	.60
.58	.58	25049	26122	27226	28361	29526	30723	31952	33213	34507	35834	.58
.56	.56	26242	27366	28523	29711	30932	32186	33473	34795	36150	37541	.56
.54	.54	27434	28610	29819	31062	32338	33649	34995	36376	37794	39247	.54
.52	.52	28627	29854	31116	32412	33744	35112	36516	37958	39437	40954	.52
.50	.50	29820	31098	32412	33763	35150	36575	38038	39540	41080	42660	.50
.48	.48	31013	32342	33708	35113	36556	38038	39560	41121	42723	44366	.48
.46	.46	32206	33586	35005	36464	37962	39501	41081	42703	44366	46073	.46
.44	.44	33398	34830	36301	37814	39368	40964	42603	44284	46010	47779	.44
.42	.42	34591	36074	37598	39165	40774	42427	44124	45866	47653	49486	.42
.40	.40	35784	37318	38894	40515	42180	43890	45646	47447	49296	51192	.40
.38	.38	36977	38562	40191	41866	43586	45353	47167	49029	50939	52898	.38
.36	.36	38170	39805	41487	43216	44992	46816	48689	50611	52582	54605	.36
.34	.34	39362	41049	42784	44567	46398	48279	50210	52192	54226	56311	.34
.32	.32	40555	42293	44080	45917	47804	49742	51732	53774	55869	58018	.32
.3	.30	41748	43537	45377	47268	49210	51205	53253	55355	57512	59724	.30
.2	.28	42941	44781	46673	48618	50616	52668	54775	56937	59155	61430	.28
.2	.26	44134	46025	47970	49969	52022	54131	56296	58518	60798	63137	.26
.24	.24	45326	47269	49266	51319	53428	55594	57818	60100	62442	64843	.24
.2	.22	46519	48513	50563	52670	54834	57057	59339	61682	64085	66550	.22
.2	.20	47712	49757	51859	54020	56240	58520	60861	63263	65728	68256	.20
.1	.18	48905	51001	53156	55371	57646	59983	62382	64845	67371	69962	.18
.1	.16	50098	52245	54452	56721	59052	61446	63904	66426	69014	71669	.16
.1	.14	51290	53489	55749	58072	60458	62909	65425	68008	70658	73375	.14
.1	.12	52483	54732	57045	59422	61864	64372	66947	69590	72301	75082	.12
.1	.10	53676	55976	58342	60773	63270	65835	68468	71171	73944	76788	.10
.0	.08	54869	57220	59628	62123	64676	67298	60990	72753	75587	78494	.08
.0	.06	56062	58464	60935	63474	66082	68761	71511	74334	77230	80201	.06
.0	.04	57254	59708	62231	64824	67488	70224	73033	75916	78874	81907	.04
.0	.02	58447	60952	63528	66175	68894	71687	74554	77497	80517	83614	.02
.0	.00	59640	62196	64824	67525	70300	73150	76076	79079	82160	85320	.00
	−1.00	119280	124392	129648	135050	140600	146300	152152	158158	164320	170640	−1.00

CORRELATION COEFFICIENTS

VALUES OF r CORRESPONDING TO VALUES OF ρ*

$$r = 2 \sin \frac{\pi}{6} \rho)$$

ρ	r	ρ	r
.00	.00	.52	.54
.02	.02	.54	.56
.04	.04	.56	.58
.06	.06	.58	.60
.08	.08	.60	.62
.10	.10	.62	.64
.12	.13	.64	.66
.14	.15	.66	.68
.16	.17	.68	.70
.18	.19	.70	.72
.20	.21	.72	.74
.22	.23	.74	.76
.24	.25	.76	.78
.26	.27	.78	.79
.28	.29	.80	.81
.30	.31	.82	.83
.32	.33	.84	.85
.34	.35	.86	.87
.36	.37	.88	.89
.38	.40	.90	.91
.40	.42	.92	.93
.42	.44	.94	.95
.44	.46	.96	.96
.46	.48	.98	.98
.48	.50	1.00	1.00
.50	.52		

*The maximum change from ρ to r is .018, a difference that may ordinarily be neglected.

CLINICAL STUDY AS A METHOD IN EXPERIMENTAL EDUCATION

By Frank N. Freeman, University of Chicago

The purpose of this paper is to present some reflections upon the relationship between group study and individual study as a method of educational investigation, to discuss some of the advantages of individual study and to present as an illustration the investigation of a case of "word blindness" or "alexia" and its treatment. The illustrative study is presented first.

Analytical Study of a Case of Alexia

The subject of this study is a girl, V. P., nine and one half years old, in the fourth grade. In the Fall of 1919, when the study was begun, she had been in school three years and had been in the hands of a tutor for one year. She was, however, unable to read and was totally unfit to do the work of the grade. She was referred to the educational laboratory for diagnosis and treatment.

Inquiry revealed the fact that the father and mother were of more than average intellectual ability. The father was a very successful business man and the mother active in the life of the community, in which she took a position of leadership. It appeared, however, that the father had had difficulty in learning to read and that he was still a very poor oral reader. The difficulty with reading also extended to the father's sister.

The child herself was slow in learning to talk. According to her mother she began to talk at seventeen months and still talked indistinctly at two and a half years of age. At the present time she talks clearly and with the average fluency.

The child had been in the University Elementary School for two years, including the first grade, and had there received the ordinary instruction in reading, which includes a considerable amount of sight reading, and in which phonetic analysis is also emphasized. In addition to this the child had been instructed for a year by a tutor and this instruction had included a very large emphasis on phonetic drill. Moreover, in her second year she had been given special help in reading by her teacher. In spite of all this intensive training, the child, when first seen, was unable to read a primer as well as is a first grade child at the end of the year.

At the outset a series of tests were made in the attempt to locate the difficulty more precisely. The child had been diagnosed by a well known oculist in the city as having "word blindness.' His prognosis was very discouraging, and he advised the child's mother that she would be unable to learn to read. It is clear that we were confronted with a case which is termed in general "word blindness" or "developmental alexia."

The first step was to discover whether there was any general intellectual defect. The child was given the Binet test. This test was carried far enough to establish the fact clearly that the child is of normal general intelligence. This conclusion was re-enforced by her general reaction, speech and manner, and also by her ability in the studies of the school which do not involve reading.

In addition to the Binet test a number of other general tests of the functions which can be supposed to be related to reading and to speech processes in general were made. It has already been remarked that the child was normal in general visual sensory capacity. In order to discover whether the difficulty in reading could be related to a general deficiency in the motor processes, the tapping test and steadiness test were made. In both of these the child's reaction was normal and the ability of the right hand was superior to that of the left. Since left-handedness and speech difficulties sometimes appear together these results had some significance as indicating lack of such deficiency. From observation the child's oral speech appeared to be moderately fluent and normally well coördinated. No difficulties were observed in the auditory speech; that is, the child is able to understand spoken language normally well and is able to reproduce from auditory presentation with normal accuracy and fullness.

As a check upon ability in general intellectual operations, as well as ability to grasp certain kinds of visual symbols, several examples in simple number combinations were presented visually. The child's reaction in these respects presented nothing abnormal.

In order to make further examination of the visual perception to determine whether the defect was due to general inability to apprehend visual symbols, the child was given the Binet Pictures Test. She responded fluently by giving descriptions of the pictures, indicating that she was able to grasp the visual objects as shown in the pictures without any difficulty. A test in the recognition of more abstract forms consisted in requiring her to match the artificial forms used in Thorndike's test. She was able to identify the forms without any difficulty. In order to extend this test she was asked to match words without giving their meaning and she was able to do this readily. She was also able to cross out the *a*'s in a pied text with average speed and accuracy.

Thus far the tests have not involved visual language in the sense of the recognition or the pronunciation of words; and in none of the tests which have been thus far described, involving visual sensation, motor processes, motor or auditory speech or visual perception, is discovered anything out of the ordinary.

In visual language a test was first given in the ability to pronounce nonsense syllables. This involves the translation of visual stimuli into sound. The child was able to make this translation and to pronounce the syllables, but did so slowly. The recognition of meaning was not involved in this test.

The next test involves the opposite process of translating sounds into visual symbols, namely spelling. In this the difficulty was somewhat greater than in the pronunciation of nonsense syllables. The simple words, *the*, *and*, and *tone* were spelled correctly and fairly readily, but there was difficulty with the words *horse* and *house* and with the word *town*. The child spelled these very slowly and there was evidence of confusion from the similarity of the different words. The child's difficulty appears, therefore, in some way connected with the translation between visual and sound symbols.

We turn now to the reading tests themselves. At the beginning of the year the subject had been given the routine reading tests in company with the other children of the school. These consisted of the Gray Oral Reading Test and the Courtis Silent Reading Test No. 2. In both cases the child was recorded as having made no score. At the time of the examination made by the writer the Gray Silent Reading Test was given. The easiest specimen, entitled "Tiny Tad," which is designed for second and third grades, was given. The rate of reading was .5 words per second and the quality was calculated as .23. The average rate for the second grade is 1.96 and the quality .252. In order to equal the quality of the second grade, therefore, the child's reading was reduced to one quarter normal speed. When given primer material she read very slowly and hesitantly and frequently became confused. It seemed useless to attempt to give other standard tests since the reading was so poor as to be below the range which is ordinarily represented in the scores of these tests.

In order to get an objective record of the child's behavior in reading photographs were taken of her eye movements. The first reading of a passage of simple material is shown

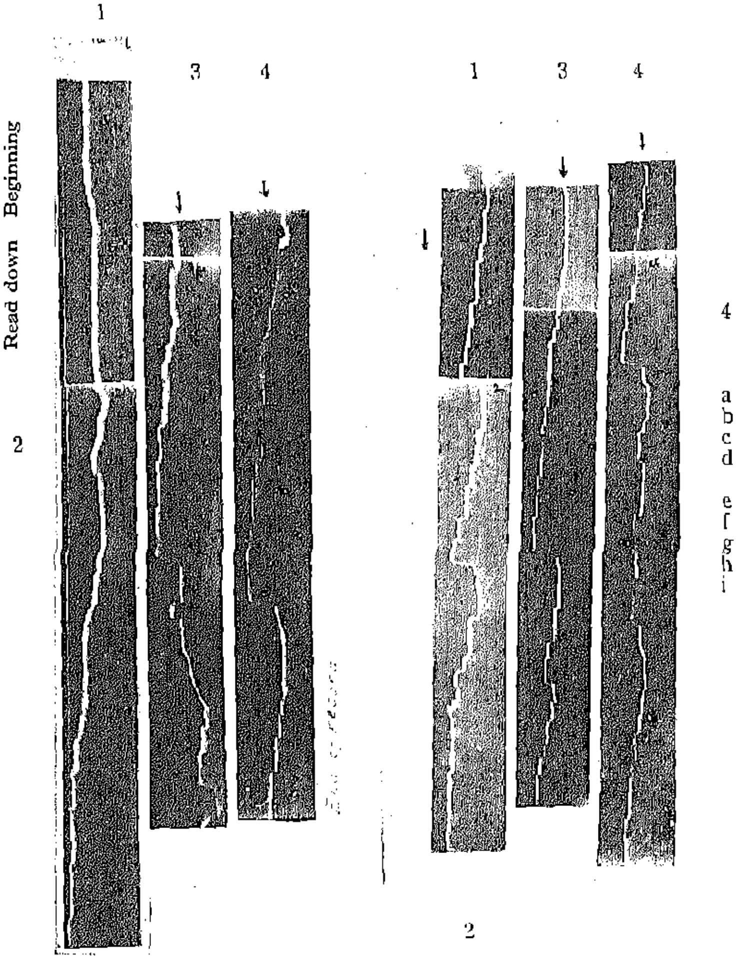


FIG. 1.—Photographic record of eye movements in reading taken at the beginning of training.

at the left of Fig. 1. In order to interpret this figure it is necessary to know what kind of record would be made in normal reading. A fairly good example of this may be seen in the record of the first line of the second reading, Fig. 1. This record consists in a series of straight vertical lines interrupted by short jogs to the left. The short vertical line represents a steady fixation of the eye and the jog to the left represents the movement of the eye along the line toward the right to the next fixation point. Reading should consist of a series of steady fixations of comparatively uniform length, interrupted by regular movements across the page. It appears from this second record that V. occasionally used this typical reading movement, but it occurred only on the second reading of the same text and even here only for a small part of the text.

Observe now in contrast the record of the first reading shown on the left of the figure. Instead of a series of fixations interrupted by abrupt movements there is here a gradual shifting or wandering of the eye. There is apparent no regular rhythmical series of forward movements and fixations, and the movement is so irregular that it is impossible to determine what positions in the record correspond to the beginning and the end of the line. From our knowledge of the ordinary type of eye movement in reading, it may be said with confidence that there is nothing here that corresponds to the typical reading habit. The irregular and uncoördinated eye movement must correspond with confused mental processes. An irregular movement of the eye indicates that the attention does not progress in an orderly fashion along the sentences and the parts of sentences.

In the second record, in which the same text was read over again, the first two lines were read in fairly well organized fashion, but in the third line confusion begins and continues throughout the passage. Observe, for example, the last line shown on the record represented on the right hand side of the figure. This line begins about one third the distance down the figure, at the point where the record line shifts toward the right. The first fixation (a) is evidently not at the right point, and after a brief space the eye shifts still further toward the beginning of the line (b). It does not remain here, however, but shifts back to the original position (c) for an interval which is less than ordinary fixation time, then moves forward and pauses (d), again moves back again to nearly its original position (e) and makes a slight shift toward the end (f). This fixation is again abandoned by a movement which apparently is not horizontal. The horizontal movements of the eye give a record which slopes somewhat downward, due to the movement of the film. Since the movement in question is represented by a horizontal line, it must have been an upward movement. The next fixation is not a steady one, but is interrupted by another backward shift. The eye then returns toward the beginning of the line (i). It is about at the point where it started from and the probability is that very little if anything has been read. Even this wholly irregular succession of movements is much more definite than the wandering movements represented in the record from the first reading.

It is obvious that the instruction in reading which the child had received up to this time had failed lamentably in the development of the characteristic eye movement habit. In the first reading of the passage, which is well within the child's comprehension, the eye movement is utterly disorganized, and even in the second reading of such a passage the irregular disorganized movement is predominant. We may conclude from these records that the movement of attention in reading is highly irregular. Instead of going forward step by step, it skips about, sometimes jumping to a point ahead of where it should be and at other times moving backward over the part which has already been read. This irregularity is in all probability due to the child's failure to grasp the meaning of the words which are fixated by the eye. Failure to grasp the meaning results in the return of the eye to the parts already fixated and in a slow wandering movement or a succession of movements made at short intervals rather than a series of clear cut movements just long enough to cover the space which can be fixated at a single pause.

On the basis of this examination the child's condition was diagnosed in the following way. There is no deficiency in general intelligence. Furthermore, the child's vision is entirely normal and she is able to recognize and to interpret the meaning of ordinary visual percepts. There is no general motor deficiency or general language disturbance. The defect, therefore, must be a highly specialized one and apparently consists in an inability to make the association between visual symbols and the sound of the words. When a child begins to learn to read he already has made the association between the word sound and the meaning. In the ordinary procedure he then proceeds to make the further association between the visual percept of the word and the sound, and by this means forms the visual-meaning association through the intermediate agency of the sound. In the case of this child the connection between the printed word and the meaning was very imperfectly made. Upon attempting to read the child began to sound the words, but did this only with great difficulty, and arrived at the meaning very slowly and very imperfectly. In this process the attention became confused and the eye movement incoördinated.

This condition was present in spite of intensive phonetic drill. The consequence of this drill was seen in the fact that the sound characteristics of words were uppermost in the child's mind. In fact they so obtruded themselves upon her attention that they formed the chief meaning of reading to her. The aspect of phonetics which preoccupied her attention was the sound of the individual letters and her reading consisted in the attempt to get the sound of the word from these abstract letter sounds. This attempt was also uppermost in her spelling. For example, she spelled the word *dyer*, *dir*. It is quite obvious how she arrived at this spelling. The *d* and the *i* are quite phonetic, and the letter *r* is commonly pronounced *er* in the teaching of phonetics. The child was, therefore, more correct than her teachers, for it is obviously inaccurate to represent the sound of the letter *r* by the syllable *er* as is regularly done. This illustrates, by the way, the pitfalls of a highly analytical type of phonetic drill.

The conclusion was arrived at that for this child, at least, phonetic drill had been carried beyond the point where it was useful. Instead of being the means to the recognition of word meaning it had become an end in itself, and really blocked the recognition of the meaning. The treatment,[1] therefore, had as its first object the short-circuiting of this roundabout association and the attempt to develop a more direct association between the sight of the words and their meaning. For the time being, then, all phonetic analysis was abandoned, but later on some attention was given to the syllabication of words in order to develop the recognition of the typical sounds represented by the different letter combinations. But this analysis was always made with real words, and the habit of recording letters as mere sounds was broken up.

As an incident to these processes an attempt was made to bring about a much more fluent word recognition. It will be remembered that the child's speed of silent reading was about one fourth of that which is normal for the second grade child. For this purpose easy reading material was given and the child was encouraged to run through the sentences in order to get the meaning rapidly. In this type of training, however, an opposite difficulty was encountered, which grew out of the fact that the child had not developed coördinated eye movements or a regular progression of attention. It was found that she was very apt to shift about on the page, to fail to read a considerable part and to fill in the gaps by guessing. It was necessary, therefore, to restrict her reading of larger units for some time and to compel her to give close attention to every part of what she read.

---

[1] The training was very efficiently conducted by Miss Claire Moore.

In the earlier stages certain devices were used to compel this attention to all the details of the passage. The passage was broken up into sentences and the individual sentences typewritten on separate slips of paper. The child then read the passage sentence by sentence. Another device was to place a card upon the page and move it forward as rapidly as the child read. In addition to these methods flash card work was given, and some use was made of printed directions which the child followed out by appropriate action. This was found to be of considerable interest to her and stimulated her to give attention to meaning. It was also found useful to give her practice in reading problems in arithmetic, which forced her to give attention to the meaning of each element in the sentence.

In addition to these drill devices the child was given continuous reading material which at the beginning was very easy. This was for the purpose of encouraging fluency without the loss of meaning. The difficulty of material was advanced as rapidly as the child could go, and a certain amount of work was also given with still more difficult material because of its inherent interest to the child. Comparatively brief periods of intensive work with difficult material was found to be stimulating and to be helpful in carrying her to a higher level of recognition than was habitual.

Parallel with this specific reading instruction there was practice in spelling and in writing words and sentences. The child was very poor indeed in spelling and her development in this respect still remains behind the development in reading. Probably this deficiency will be permanent.

Along with the practice in reading itself there was a good deal of work which was intended to direct the child's attention to the meaning of what was read, to prepare her to grasp the meaning by anticipating the difficulties. Before any passage was read, the general topic was discussed and the child's curiosity in it was aroused. At the same time particularly difficult words were written on the board and studied. In this way the habit of slurring over unknown words or of pausing too long to study them was avoided.

The period of experimentation which is covered by this report was from the early part of October until a little before Christmas. Two weeks were lost from this period by sickness. The child was in the regular school grade except for forty-five minutes five days in the week, when she was given the type of training which has been described. At the end of this period reading tests were again given and another photograph was taken of the eye movements.

It will be remembered that the child made no score in the Gray Oral Reading Test at the beginning of the year. After about ten weeks of training she made a score of 36.25. This is about half way between the norms for grades I and II. In the Courtis Silent Reading Test, she read at a rate of 3.2 words per second, which is six times as fast as at the beginning and equals the standard for grade VI. This was too rapid in comparison with the child's understanding, and indicates that it was necessary to check her speed and to emphasize the meaning. This has accordingly been done in subsequent training. The number of questions answered in this test was 34, which is half way between the standards for grades IV and V. The index of comprehension was .62, which is between the standards for grades II and III. This index, however, is somewhat unfair to the child. An examination of her paper showed that she answered the questions much more accurately during the first three minutes of the test, but that then she lost control and began to guess. This indicates the instability of her reading habit. If we take the first three minutes we find that the number of questions was 32, equivalent to the standard for grade IV, and the index of comprehension .79, which was equal to the standard of grade III. Thus the child has made progress in the two weeks training equivalent to perhaps three years ordinary progress in a school.

At this same time a second photograph was taken of the child's eye movements. This photograph is shown in figure 2. A different passage was used from that used in making the first photograph. While some of the record is dim, it is clear enough to indicate great progress in the development of a well organized eye movement. No line was read in as disorganized a fashion as was typical of the earlier reading. Most of the lines were read with a fairly well organized eye

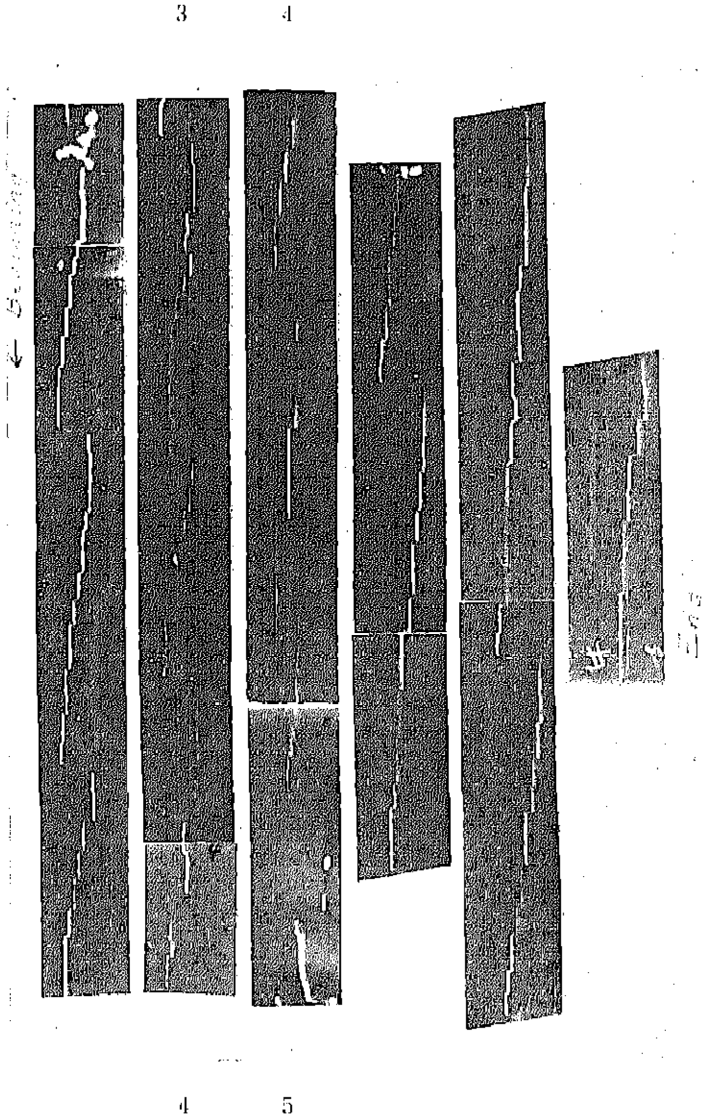


FIG. 2.—Photographic record of eye movements in reading taken after ten weeks' training.

movement habit. The progress in the eye movement habit thus runs parallel to the progress in reading ability.

The prediction of the oculist that V. would not be able to learn to read has been proven to be unfounded. Her mother reports that she shows spontaneity and interest in reading at home. She is still considerably below the standard for her grade and probably will always be a reader of very moderate ability. There is no evidence, however, that she need suffer the serious handicap of illiteracy. The training of the child is now being continued and in order to make her work harmonious she is being given special instruction in other subjects in connection with the work in reading. The hope is that eventually she may return to the regular school grade.

Conclusions from the Individual Study

From this case we seem warranted in drawing the following conclusions.

1. We seem justified in raising a serious question whether there is such a thing as specific congenital word blindness or "alexia." It is doubtful whether there are any children of normal intelligence and vision who are unable to learn to read. This conclusion is supported by the experiments of Dr. Clara Schmitt,[2] and Supt. C. J. Anderson and Miss Merton.[3]

2. In treating a case which presents special difficulty it may not suffice to simply give more intensive drill of the kind which is suited to average children. It may often be necessary, and perhaps is usually necessary, to give a specially adapted drill. In any case it is necessary to follow the development of the individual case, to examine the child to find out just what his stage of progress and his difficulties are, and to adapt the training moment by moment to the stage at that particular time. Supt. Anderson's work indicates that this type of diagnostic treatment can be carried on by the room teacher, and does not necessitate special laboratory facilities. It will, of course, require a somewhat unusually trained teacher to carry on such work at the beginning.

3. There are certain limitations to the usefulness of a conscious phonetic analysis in reading. There are two possible interpretations of the case. The first is that there are certain individuals for whom conscious analysis is more difficult than

---

[2] Clara Schmitt, Developmental Alexia. *Elementary School Journal*, 1918, vol. 18, pp. 680, 757.

[3] C. J. Anderson and Elda Merton, Remedial Work in Reading, *Elementary School Journal*, 1920, vol. XX, pp. 685-701 and 772-791.

for others, and with whom this analysis should be very much reduced if not eliminated. The other possibility is that this conscious analysis is suited only to certain periods in the development of reading and that if it is carried beyond this period it greatly retards progress. Both of these suppositions are possible in the present case. Which ever one is correct we may at least make the general statement that phonetic analysis when made consciously has limitations which should not be overlooked.

4. The lack of proper development of word recognition is accompanied by the formation of faulty eye movement and attention habits. Special practice is required to overcome these faulty habits when they have been developed through ill adapted previous training. It is possible for special training to bring about radical improvement in the eye movement.

5. The extension of the method of individual diagnosis and training promises very large economy of time and effort in the work of the school. How far we should go in the direction of individual instruction no one can predict. There are good reasons to believe that we should not abandon class instruction. If we are to retain it, however, we must find some means of overcoming the tremendous waste which results from a number of children marking time as a consequence of the lack of adaptation of the work to their individual requirements.

The need of individual attention in the school is paralleled by the need of individual study as a method of research. The rest of the article is a comparison of the individual and the group methods.

The Need of Analytical Study of Individuals

The prevailing method of study in education, and to perhaps a lesser degree in psychology, is the investigation of groups. Experimental psychology began with an intensive and analytical study of individuals. The subject of individual psychology or of individual differences was scarcely considered in the early period. Individual differences were in fact looked upon implicitly, if not with full consciousness, as errors to be disregarded, or to be overcome by a more perfect technique. With the direction of attention upon individual differences following the work of Stern, Binet and Galton, there has come into use a much more extensive and comparative study of groups and as an instrument to this study there has been developed an elaborate statistical technique. This statistical technique has enabled us to discover facts and to make statements about the group which previously we were not able to make. There is no reason for undervaluing the importance of this method and the light which its use has thrown upon facts relating to groups. There are some questions which the group method is capable of solving. There are other questions, however, for the study of which an individual analytical study is equally if not better suited, and it is desirable, in view of the current disposition to neglect it in comparison with the group method, to analyse the situations in which the individual method is applicable.

Group study is applicable to at least three types of problems. First, the determination of central tendencies, second the determination of variations from the central tendencies, and third the discovery of relationships. We have only to glance at the studies in the field of psychology during the first twenty or twenty-five years of its development, to find a plentiful number of cases in which the crudest types of calculations were made. Statements in regard to the average performance of the group were made without regard to the most elementary necessities of statistical method. The number of cases might be too few, or type of distribution might be disregarded, or the question whether the average or the median was the best measure of central tendency to use was commonly disregarded. As a consequence the question of the reliability of such statements was not raised. This was particularly pernicious when the averages of two groups were compared. Frequently the interchange of a single case from one group to another might alter the averages sufficiently to change the relative position of the two groups. Furthermore, the range of the variation of the individuals within a group, the character of the variation, whether according to the normal frequency distribution or some other type, the relation of the variation to a comparison of averages, were not treated at all in many, if not most of the early studies. These elementary principles regarding the statistical treatment of measures of the group are now to be taken for granted.

When we come to the study of relationships, however, the honors are divided between the study of large groups and the study of individuals. There are many problems raised by means of the study of groups which can be settled only by the addition of the study of individuals. A few examples may be cited of problems which have been studied by the group method but in which the results are not entirely satisfactory and in a solution of which it would be very desirable to add to the group study the study of individuals.

Consider, for example, the various methods of learning. The comparison between the efficiency of the part method and the whole method of memorizing has been investigated by having a group memorize first by one method and then by another, or, by having one group use one method and another the second method. The results of learning by the part method and the whole method are then averaged and from the comparison of these averages the conclusion as to which is the better method is drawn. Now such a conclusion might be statistically sound and yet psychologically false. It might, for example, be true that over half of the individuals selected at random could learn better by the whole method and fewer than half, say forty per cent, could learn better by the part method. If a large number of individuals are given both methods it may be an accurate statement in reference to the group to say that on an average the whole method gives better results because in the long run the greater number of individuals can use this method more advantageously. However, it might also be true that the forty per cent of individuals were so constituted that they could always learn better by the part method. The average results, therefore, could not apply to them.

Take another illustration. There have been recent attempts made to discover what is called the general form of the practice curve. This attempt may consist in putting together, by means of the average, the results of the practice curves of different individuals, or of putting together practice curves of different types of learning. The second method is perhaps not so readily accepted as the first, but even this, in the judgment of the writer, is not justified. The only assumption on which it is justified is that the variations of the curves which represent the progress of different individuals are accidental. In other words, if one individual gains more rapidly at the beginning, and another one more slowly, than the average, this difference would on this assumption be ascribed to some accident in their previous history or in the circumstances surrounding their work, rather than to a fundamental difference in their mode of learning. Certainly, at least, the burden of truth is upon this assumption; and unless it can be substantiated we must regard an average of the practice curves of different individuals as a mathematical fiction.

Again, consider the problem of the relationship between speed and accuracy. There have been a considerable number of studies, particularly in the school subjects, which have compared the relationship between the speed and accuracy of the performance of different individuals. This may be done by grouping individuals' scores according to the double criterion and arranging the groups in a single scatter or distribution table. If the individuals are arranged into three groups on the basis each of speed and accuracy, there would be nine groups in all which would be formed. By a comparison of the number of cases in the nine groups it is possible to determine whether or not individuals with high speed are likely to excel in accuracy or whether the reverse is the case. In general, the conclusion has been reached that, contrary to a widespread popular opinion, high degree of speed is likely to be associated with high degree of accuracy. This is a very valuable bit of information and enables us to predict within wide limits of variation the probability of various combinations of speed and accuracy in an individual. There has been, however, a tendency to make application of this fact to the training of individuals, which is to say the least uncertain, and which could be very much more securely founded upon an experimental and analytical study of the individuals themselves. Because an individual who works at high speed also works with superior accuracy we cannot conclude safely, as has sometimes been done, that for an individual to increase his speed will result in an increase in his accuracy. This may or may not be true, but we can only determine its truth and the limits within which such a relationship holds, by training individuals to work at various speeds and testing the effect upon the accuracy of their performance.

This leads us to the treacherous question of correlation. What does a correlation between two traits mean? Assume a correlation of .60 and assume further that the correlation coefficient expresses the real relationship. What does such a coefficient mean with reference to the relationship between the traits in question in the various individuals of the group. Obviously it may mean different things. We may say that there is a strong tendency for the two traits to be related, but we are here speaking in figurative terms. In any given individual the traits are either related closely or they are related remotely or not at all. A moderate degree of correlation might mean that there is a fairly close relationship between two traits in every individual of the group, or it might mean that the relationship was very close in some individuals and very slight in the case of others. In order to make entirely clear to ourselves what such a co-efficient means it is necessary to get within our mental grasp the varieties of the types of cases which exist in the individuals of a group. The classification of individuals into such types is necessary before we are able to determine the administrative application to make of a given degree of correlation. For example, it was found in the study of the correlation between pupils' standings in High School subjects, that there are all types of cases from those of one extreme in which the standing in the different subjects was highly consistent to those in which there was an extreme divergence in the standing in the different subjects. Intermediate between these were pupils in whom the standings in different subjects varied, but not to an extreme degree. The combination of the relation between the scores of all the pupils into a single correlation coefficient in the case of a pair of subjects would result in a moderate degree of correlation, but obviously it would not be just to treat a pupil in whom ability in two subjects was equal on the assumption that the relationship was only moderate; and furthermore, on the other hand, it would not be correct to treat a pupil in whom ability in the two subjects was opposed on the assumption that they were related to a moderate degree.

We should keep in mind continually that it is individuals with whom we are to deal and that our conclusions should always be in terms of individuals. An individual is a complex of many traits and abilities. When the question of failure in learning, for example, arises in the case of an individual, it is necessary not only to know that a condition may arise in such a certain percentage of cases from one cause, and in another percentage from another cause, etc., but it is necessary to know what condition or conditions actually produce the failure on the part of the individual before us. In such an attempt to diagnose the condition of an individual we see the necessity of the combination of the method of group study and of individual study. We proceed to investigate the ability of an individual in various processes, but the only ground on which we can determine whether this ability is high or low is by a comparison with the ability of the group to which the individual belongs. Thus in order to diagnose reading ability, arithmetic ability or spelling ability of a fourth grade child we should compare the scores which he makes with those of the norms for his age or grade. In this way group study and individual study are combined.

SIGNIFICANT RESPONSES IN CERTAIN MEMORY TESTS[1]

By JOSEPHINE CURTIS FOSTER, Formerly Chief Psychologist, Psychopathic Hospital, Boston

The psychologist who gives a large number of "mental tests" to a variety of patients comes to recognize certain responses as unusual. Perhaps, judging from failures of a subject in early tests, he gets "set" for a certain grade of answers and so is surprised at a correct response. Thus frequently an examiner will report: "There is something queer about that examination. I don't believe that patient is feebleminded, in spite of the low rating."

We have been trying at this hospital for some time to get a numerical statement of this "queerness." We have already reported our method of calculating the irregularity[2] of the Point Scale examination. We wish now to suggest other features in the psychological examination which should lead the examiner immediately to suspect that he is dealing with a psychosis or with some deteriorating process. We do not, of course, wish to claim that the psychologist should be capable of making a psychiatric diagnosis. Far from it! But we believe that the psychologist should be able to recognize certain cases as cases properly to be referred to a psychiatrist, and that he should be able to make a guess as to whether low mental age is attained as a result of lack of development or as the effect of some retrograde process in a subject whose mentality was previously normal.

What we need is not so much new tests, as the understanding and interpretation of the results from the tests which are at hand. In the present paper we shall consider the different types of reactions found in the tests of memory. There are three kinds of memory tests in the routine psychological examination commonly given at this hospital: (1) memory span for digits and sentences; (2) memory for the two Binet drawings after an exposure of 15 sec.; and (3) the recall of short paragraphs: one in year X of the Stanford revision of the Binet scale, and two Healy tests, in which the subject is required to recite what he can of a selection read aloud, in the one case by himself, in the other by the examiner.

---

[1] The writer is greatly indebted to Miss Katherine T. Gere, formerly, psychological interne at the Psychopathic Hospital, for her work in collecting data and assisting in the arrangement of results.

[2] Josephine N. Curtis, "Point Scale Examinations on the High-Grade Feeble-Minded and the Insane," *Journal of Abnormal Psychology*, XIII, 1918, p. 88 ff.

All cases considered in this paper were examined during the last two years. All but approximately 5% of the examinations have been corrected by the same person before a final score was given, thus avoiding individual variations in the grading of the separate tests. No case in which there was language difficulty, poor co-operation, or disagreement as to diagnosis has been considered. The work of computation has extended over eight or nine months. The total number of cases discussed under the different sections, therefore, is not always the same. Since the selection has always been on a uniform basis, however, the results are comparable. We shall take up the separate tests in the order mentioned, and try to show what reactions should be noted by the examiner and reported as of significance for diagnosis.

Memory Span

In this part of the investigation we used the records of patients described in Table 1.

TABLE 1

CASES USED IN DISCUSSION OF MEMORY SPAN

Number of Cases	Diagnoses Made by Hospital Psychiatrists	Average Point Scale Score	Average Chronological Age
41	"Not Insane"	83.8	29.9
46	Feeble-Minded	61.7	22.4
56	Dementia Praecox	73.9	31.2
45	Alcoholic Psychoses	70.9	41.4
16	Manic-Depressive Psychoses	75.4	44.5
13	General Paresis	58.0	35.9
10	Unclassified Paranoid Condition	79.1	42.7

We have also partial records[3] of 503 normal school children who were examined in obtaining the first norms for the Point Scale, but their average chronological age and their drawings were not available. We include their results in some tables for comparative purposes.

---

[3] These data were very kindly supplied by Dr. R. M. Yerkes.

From the calculation of these data according to diagnosis, mental age, and chronological age of the patient, we have concluded that the memory span for digits is of little significance in our present problem. To be sure, with all patients, whatever the diagnosis, memory span increases with advancing mental age; but the increase is so slight that after mental age 10 we can show it only by keeping a decimal place. In general, we may say not only that for all mental ages over 10 we may expect a memory span of six or seven digits, but also that one of only five digits is not unusual. When the cases with different diagnoses are grouped by chronological ages, again we find no significant differences. For example, for the four decades, from ten to forty-nine, our "not insane" cases give averages of 6.4, 6.5, 6.6, and 6.3 while dementia praecox patients give 5.3, 6.0, 5.3, and 5.7. Similarly patients suffering from other diseases show no regular change in the memory span with advancing chronological age. That the diagnosis itself has little or no relation to the memory span may be seen from the fact that if we use only those cases where the mental age is 11.5 or over, so that we surely get no influence of mental age, we find the average number of digits given to be: subjects suffering from dementia praecox, 5.8; from alcoholic psychoses, 5.9; from unclassified paranoid condition, 5.9; from manic-depressive psychoses, 6.1; those "not insane," 6.5; and the feeble-minded, 6.8. The greatest difference, 1.0, is too small to be considered significant in view of the small number of cases and the possible selective factors involved.

The memory span for sentences, at least as given in the Point Scale, is useless for our purposes. Practically no patient succeeded in repeating the fourth (that is to say, the longest) sentence, and practically no patient failed to repeat the two short sentences. There was, then, only the question of whether some patients could give the third (next to the longest) sentence. The scores here showed no uniformity and we concluded that the performance is poorly constructed to show a relation to mental age, to chronological age, or to diagnosis.

The examiner, therefore, need pay little attention to either of the memory span tests so far as diagnostic hints are concerned.

Memory for Drawings

In this part of the investigation we used the records of patients grouped in Table 2.

TABLE 2

CASES USED IN DISCUSSION OF MEMORY DRAWINGS

Number of Cases	Diagnosis	Average Point Scale Score	Average Chronological Age
72	" Not Insane "	80.0	18.8
183	Feeble-Minded	59.7	21.5
121	Dementia Praecox	62.7	32.1
23	Syphilitic Psychoses	64.3	41.5
14	Unclassified Paranoid Condition	73.7	48.0
42	Acute Alcoholic Psychoses	75.4	38.5
20	Deteriorating Alcoholic Psychoses	76.2	49.6
35	Manic-Depressive Psychoses	72.1	44.1
13	Arterio-Sclerotic Psychoses	55.0	61.3
10	Senile Dementia	54.6	68.3

It is evident from our results that the psychological examiner must take a number of factors into account in making an estimate of the significance of failures or successes in this test. When we fractionate our cases for, and distribute according to mental age we find a tendency for the score on the memory drawings to increase with advancing mental age. The results for patients with those diagnoses which show fairly large numbers of cases are given in Table 3. No figure is based upon less than five cases.

TABLE 3

AVERAGE SCORE ON MEMORY DRAWINGS FOR DIFFERENT DIAGNOSES AT DIFFERENT MENTAL AGES

(HIGHEST POSSIBLE SCORE IS 4)

Diagnosis	Mental Age						
	7–8	9–10	11–12	13–14	15–16	17–18	18+
Normal Children	.5	1.1	1.8	2.2	2.7	3.0	3.4
" Not Insane "		.5	1.9	2.8	2.8	3.0	3.3
Feeble-Minded	.8	1.8	2.1	2.7			
Dementia Praecox		.5	.9	1.8	2.1	1.6	3.0
Acute Alcoholic Psych.		.4	.8	.7	1.4		2.3
Manic-Depressive		.6	1.2	1.0			1.5
All Hospital Cases	.5	1.2	1.3	1.8	2.3	1.8	2.7

From Table 3 it appears that the insane are well below normal children, "not insane," and feeble-minded of the same mental ages, but that scores tend to increase with advancing mental age.

When we group our cases by chronological ages we find, on the contrary, that the average score tends to drop off with advancing age, at least after age 40. The results are given in Table 4.

TABLE 4

AVERAGE SCORE ON MEMORY DRAWINGS FOR DIFFERENT DIAGNOSES AT DIFFERENT CHRONOLOGICAL AGES

(HIGHEST POSSIBLE SCORE IS 4)

Diagnosis	Chronological Age						
	10–19	20–29	30–39	40–49	50–59	60–69	70–79
" Not Insane "	2.9	2.9					
Feeble-Minded	2.2	1.6	1.1				
Dementia Praecox	1.4	1.6	2.0	.9	.7	.0	
Syphilitic Psychoses		.9	.0	.6			
Deteriorating Alcoholic				1.8	.3		
Acute Alcoholic		1.6	1.8	.9			
Manic-Depressive		2.0	1.2	2.0	.6	.5	
All Hospital Cases	2.3	1.7	1.4	1.1	.6	.2	.1

From the two tables just given we see that although cases with all diagnoses show increase in score on memory drawings with increase in mental age, and decrease in score with increase in chronological age after age twenty or thirty, there are, nevertheless, decided differences from diagnosis to diagnosis as to the exact age at which the change begins or is most marked. When we come to the question of the disease itself, it is therefore necessary to reduce as far as possible the influence of mental and chronological age. In order to do this we have discarded those cases with mental ages below 11.5. This means that we have also eliminated practically all cases with chronological ages of 50 or more, for none of our older patients received high scores on the Point Scale. With such limitations we find the average scores on the test to be: cases diagnosed as "not insane," 2.9; normal children, 2.7; cases diagnosed as feeble-minded, 2.7; as dementia praecox, 2.1; as manic-depressive, 1.5; as acute alcoholic psychoses, 1.5; as deteriorating alcoholic psychoses, 1.4; as syphilitic psychoses, 1.2; as unclassified paranoid condition, .8; and of those diagnosed as arterio-sclerotic psychoses and senile dementia, there were too few cases. If further restriction is made by discarding all cases with chronological ages of over 30 we get the same order of diagnoses for the first six groups. The other groups drop out. The mean variations for these averages are all about .9.

In the score for the drawings from memory, therefore, we find the first point at which the psychological examiner should be on the lookout for inconsistencies. Realizing that a high score is to be expected from young persons and from persons with a high mental age, the examiner should note any case where low chronological age and high mental age are combined with poor performance in the memory drawings.

Not only is the score on the memory drawings important, but the shape and size of the drawing itself may be indicative of mental disease. A drawing which receives no credit may be as significant as one which adds to the total score. If we take the drawings which were not given any credit, we may arrange them in the following groups: those giving the correct drawing, but adding extra lines (e. g., in the first, the "box," drawing, when the diagonals were continued to meet within the inner rectangle, or in which there was more than one inner rectangle, or in the second drawing when two or more humps were made between the uprights); those giving correct drawings except that they omitted a few lines; those giving very sketchy drawings (e. g., in the first, drawing merely one rectangle, in the second, drawing merely a jointed line); those making drawings which showed they had the general idea only (e. g., in the first, making a conventional box, or a rectangle with something inside, in the second, a horizontal line with curlicues or ornaments of some kind at the ends); those drawing things which showed practically no resemblance whatever to the original (e. g., a triangle for the first, or a circular figure for the second); and those making no attempt to draw the figures at all. The results we obtained from this grouping were significant but were so awkward to handle that we combined the groups into those giving the general idea (as above), those giving sketchy drawings (including also those who made no attempt, and those who omitted lines), and those giving elaborated drawings (including also those whose drawings showed no resemblance to the original). Our data give no evidence that mental age and chronological age have any influence on the type of memory drawing. The cases with the different diagnoses, however, give rather typical results. When we consider only those cases with mental ages of 11.5 or more (so that any influence of mental or chronological age is even more certainly eliminated) we find those inclined to elaborate or fabricate in the drawings to be those suffering from syphilitic psychoses (43% elaborated *vs* 0% who give scanty drawings); from dementia praecox (24% elaborated *vs* 12% scanty); and from unclassified paranoid condition (40% *vs* 30%). On the other hand, those tending to give scanty or sketchy drawings are: those suffering from deteriorating alcoholic psychoses (33% scanty *vs* 0% elaborated); and the feeble-minded (19% scanty *vs* 5% elaborated). The cases diagnosed as "not insane" give less than 10% in either scanty or elaborated drawings. Illustrations of such drawings are given in Plate 1. Figure (a) shows a sketchy drawing of the "box" made by a case of senile dementia, chronological age 75, mental age 7.5. Figure (b) shows a slightly elaborated drawing of the first figure made by a case of dementia praecox, chronological age 22, mental age 11.2. Figure (c) shows practically no resemblance to the original "box" which the patient, a case of general paresis, chronological age 50, mental age 8.9, was attempting to reproduce. The other figures in the plate are reproductions of the second memory drawing. Figure (d) was made by a case of manic-depressive insanity, chronological age 30, mental age 18.0 and belongs, of course, to the "scanty" type. Figures (e) and (f) are examples of elaboration and were given by a case of dementia praecox, chronological age 18, mental age 13.5 and by a case of general paresis, chronological age 30, mental age 11.6. Figure (g) shows an entirely fanciful reproduction by a case of general paresis, chronological age 41, mental age 8.9.

Our main conclusions considering all computation of results on both drawings, individually and together, cases with mental ages 11.5 or over, and all mental ages together, types of drawings as originally separated, and as grouped together, are these: Patients suffering from syphilitic psychoses, dementia praecox, and unclassified paranoid condition tend to elaborate the original drawing, or to make some fanciful combination of lines having little or no relation to the original. Cases of arterio-sclerotic psychoses and senile dementia are apt to fail to attempt the drawings at all. Cases of manic-depressive insanity, arterio-sclerotic psychoses and senile dementia are apt to omit lines. The feeble-minded omit lines more often than they add them.

The size of the drawing from memory is probably also of significance. Patients are not instructed to reproduce the drawings in any special size, though the size is roughly limited by the dimensions of the record sheet (10.5 in. by 7.5 in.). If a patient asks "Shall I make it just the same size as the one I saw?" he is told "Why, yes. Just about. You need not be too fussy." Very few patients ask. We shall confine ourselves to the discussion of the first drawing, since it is much the easier to measure, and we have observed that increase or decrease in the one drawing tends to carry over to another. The height of the original drawing is 4.4 cm. Our records give no evidence that mental and chronological age have any influence on the size of the drawings from memory. The average heights of the drawings given by patients are: those diagnosed as "not insane," the feeble-minded, and those suffering from deteriorating alcoholic psychoses, each 3.8; those suffering from manic-depressive insanity, 4.1; from acute alcoholic psychoses, and from senile dementia, each 4.2; from arterio-sclerotic psychoses, and from unclassified paranoid condition, each 4.3; from demential praecox, 4.4; and from syphilitic psychoses, 5.3. (The mean variations of all these averages are about 1.0.) Similarly, if we consider the percentage of cases giving extremely large or extremely small drawings, we find that the cases of syphilitic psychoses, of unclassified paranoid condition, and of dementia praecox tend to exaggerate the size; while the feeble-minded and the cases of manic-depressive and alcoholic psychoses tend to diminish the size.[4] These groups are the same that we found in the discussion of the "types" of the drawings. Thus, roughly speaking, paucity of detail and relative smallness of size are found in one group of diagnoses, elaboration of detail and largeness of size in another group.[5]

We may summarize our discussion of the drawings from memory as follows: The psychologist should note: (1) a low score on the drawings from memory combined with a comparatively low chronological age and a high mental age; (2) elaborated or fanciful drawings; and (3) very large drawings.

---

[4] In 370 examinations of rather low-grade normals which we have consulted in connection with another problem, the average height of the drawings was 3.3 ± .8.

[5] It would be interesting if we could know how many of these patients show enlarged handwriting under distraction, as in the work reported by June E. Downey on the "Will Profile," University of Wyoming Bulletin, Vol. XV, No. 6A, 1919.

Memory for Short Paragraphs

In this part of the investigation we used the same records as in the discussion of the memory span.

It appears from our results that the tests now to be considered are highly significant.[6] The results for patients suffering from the various diseases at different mental ages where there were a fairly large number of cases are given in Table 5. In this table we have averaged the numbers of details given in the three selections for the sake of simplicity and because there were no obvious variations in the results from selection to selection.

TABLE 5

AVERAGE NUMBER OF DETAILS RECALLED IN PARAGRAPH MEMORY TESTS BY DIFFERENT DIAGNOSES AT DIFFERENT MENTAL AGES

(HIGHEST POSSIBLE AVERAGE IS 17.3)

Diagnosis	Mental Age						
	7–8	9–10	11–12	13–14	15–16	17–18	18+
"Not Insane"			8.7	12.4	9.2		12.3
Feeble-Minded	3.4	6.7	8.4	11.6			
Dementia Praecox	2.1	4.7	6.2	7.0	8.6	9.6	10.0
Alcoholic Psychoses	3.3	4.6	7.0		8.7	9.5	8.0
Manic-Depressive	3.4	8.0	6.0			8.9	
Unclassified Paranoid			6.2	5.0			9.7
General Paresis		6.6					8.4
All Hospital Cases	2.7	5.7	7.3	8.2	8.9	9.8	9.8

In general, as in the drawings from memory, we find for each disease, as well as for all diseases grouped together, a tendency for the number of details to increase with an increase in mental age. We see, moreover, that the cases diagnosed as "not insane" are decidedly superior to the insane of the same mental age. This apparently means that memory deteriorates or shows irregularities more quickly than intelligence at large.

When we group the cases by diagnoses and by chronological ages, we have Table 6.

---

[6] Doubtless the test would be still more useful if each part were read to the subject, instead of having him read some of them aloud himself. In all work with psychotic persons we find many patients with poor eyesight, and with scanty education. It seems clear that any difficulty with reading thus caused may constitute an unfair distraction and lower the score.

TABLE 6

AVERAGE NUMBER OF DETAILS RECALLED IN PARAGRAPH MEMORY TESTS BY DIFFERENT DIAGNOSES AT DIFFERENT CHRONOLOGICAL AGES

(HIGHEST POSSIBLE AVERAGE IS 17.3)

Diagnosis	Chronological Age				
	10–19	20–29	30–39	40–49	50–59
"Not Insane"	11.5	9.8	9.8	8.4	
Feeble-Minded	9.0	7.5	5.6	4.5	
Dementia Praecox	7.5	6.5	7.2	6.5	
Alcoholic Psychoses			7.4	7.3	6.4
All Hospital Cases	9.7	7.2	7.9	7.0	7.1

In general we find here, as in the case of the drawings from memory, that the score in the paragraph memory test falls off with advancing chronological age. This is particularly marked in the case of the feeble-minded.

Again, as in the earlier tests, we have endeavored to reduce as far as possible the influence of mental and chronological age by considering only those cases with a mental age of 11.5 or more. When we do this we find the average number of details recalled to be: by the "not insane" cases, 10.1; by the feeble-minded, 10.0; by those suffering from manic-depressive psychoses, 8.6; from dementia praecox and alcoholic psychoses, each 8.4; and from unclassified paranoid condition, 7.8. We have here evidence of a tendency for the feeble-minded and the "not insane" cases to give more details than the psychotic patients of the same mental level. It will be remembered that by limiting our mental ages we limited the chronological ages to those under 50, but it is evident from Table 6 that the reduction in score with advancing chronological age begins long before 50 is reached. If we reduce still further the influence of chronological age by considering, of our group over mental age 11.5, only those with a chronological age under 30, we deplete the ranks to such an extent that only three groups remain. The average scores for these are: "not insane" cases, 10.7; the feeble-minded, 9.9; and patients suffering from dementia praecox, 7.9.

We have shown that dementia praecox patients, and probably other psychotic cases, give fewer details in the memory tests than do the "not insane" cases and the feeble-minded of the same mental age. We come now to the question of the number of errors made. By errors we mean such responses as "December" for "September," "tenement" for "house," etc. We find such errors for all ages and for all diagnoses.

The number of errors decreases slightly with advancing mental age, and apparently does not change with chronological age. Errors are least frequent among the "not insane" subjects, and most frequent among the dementia praecox and general paretic patients. Indeed, nearly half that the general paretics give is wrong.

If we disregard such simple mistakes and consider merely the cases where the subject adds something which was not in the original, we find all except the feeble-minded making additions of some sort. In the case of the "not insane" the number is very small and all additions have a close connection with the material which had been read. Some of the patients, however, give irrelevant material.[7] Of the general paresis cases who had a mental age of 11.5 or more, 70 per cent gave additional or irrelevant details. Similarly 27 per cent of the cases of manic-depressive insanity (and all of these 27 per cent were manic), 16 per cent of those suffering from alcoholic psychoses, and 4 per cent of dementia praecox patients give additional details. This list agrees for the most part with the list of those adding to the drawings from memory. In the drawings we found the patients most likely to add to be those suffering from syphilitic psychoses (corresponding to the general paresis cases here), from unclassified paranoid condition (not represented here), from dementia praecox (as here), and from acute alcoholic psychoses (corresponding to the alcoholic psychoses here).

In the memory for short paragraphs, therefore, as in the memory for drawings, the psychologist must watch for inconsistencies between the mental and chronological ages and the score on the test, and for any cases which add details, particularly those cases which add irrelevant matter.

Conclusions

1. There is no evidence that the insane give responses in the memory span for digits or sentences which are markedly

---

[7] Illustrations of what we mean by "irrelevant material" are given by two records, each taken from the examination of cases of general paresis. One reproduced the Healy Visual Verbal test beginning "If a man . . ." as follows: "The girl was burnt bad and he went and got firemen and he was burnt. If he didn't report it, people who were there would have been burnt too." The other patient reproduced the Healy Auditory Verbal selection beginning "If a sailor . . ." as follows: "There was a boy going to a ship, was always looking for water. Then he was looking for a cargo and looking out for trains going across."

different from those given by normal persons of the same mental and chronological ages.

2. The responses given by the insane in the tests for memory drawings and memory for short paragraphs are less adequate than those given by the normal and the feeble-minded of the same mental and chronological ages. This is probably due to the fact that many of our insane cases suffer a deteriorating or temporarily incapacitating process which, like increasing chronological age, affects memory to a greater extent than it does intelligence at large.

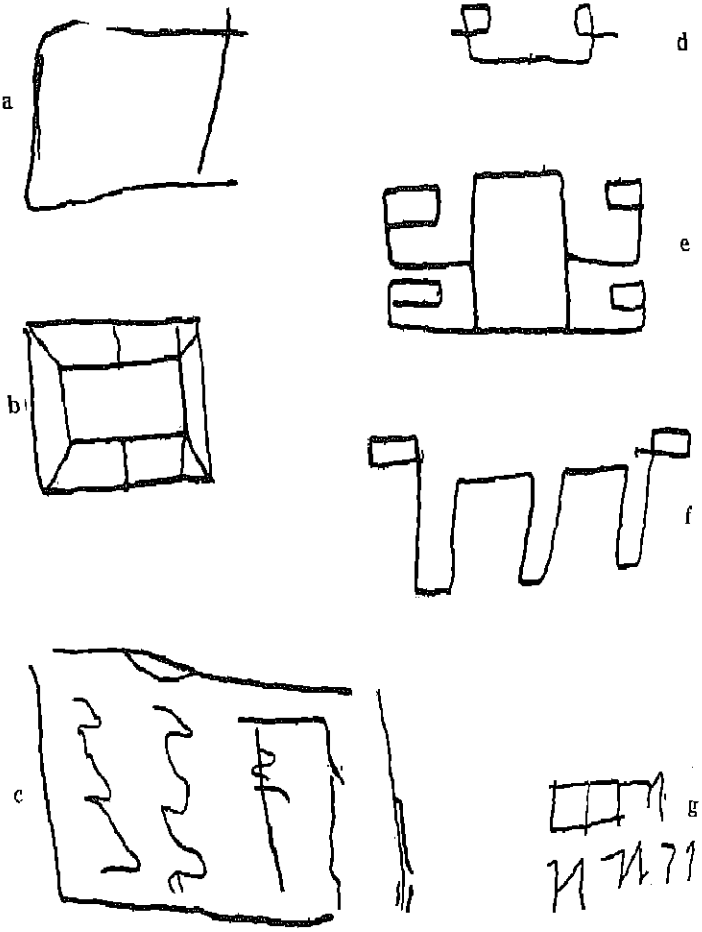
3. In certain types of insanity there are significant responses not indicated in the score, such as extreme size and irrelevant details in the case of the memory drawings, and numerous errors and additions in the reports of the paragraph selections.

4. Table 7 summarizes some of the chief types of response which may be expected and which may be given some weight as diagnostic signs, for cases with high mental ages (at least 11.5) and comparatively low chronological age (under 50).

TABLE 7

RESPONSES TO BE EXPECTED FROM DIFFERENT DIAGNOSES

Diagnosis	Drawings			Paragraphs	
	Score	Size in cm.	Type	Number of Correct Details	Irrelevant Additions
"Not Insane"	2.9	3.8	Correct	10.1	None
Feeble-Minded	2.7	3.8	Correct	10.0	None
Dementia-Praecox	2.1	4.4	Elaborated	8.4	Few
Syphilitic Psychoses	1.2	5.3	Elaborated	(few)	Many
Alcoholic Psychoses	1.5	4.0	Sketchy	8.4	Few
Manic-Depressive	1.5	4.1	Sketchy	8.6	Some
Unclassified Paranoid	.8	4.3	Elaborated	7.8	Few



(*a*) Senile dementia, chr. age, 75; mental age, 7.5.
(*b*) Dementia praecox, chr. age, 22; mental age, 11.2.
(*c*) General paresis, chr. age, 50; mental age, 8.9.
(*d*) Manic-depressive (depressed), chr. age, 30; mental age, 18.0.
(*e*) Dementia praecox, chr. age, 18; mental age, 13.5.
(*f*) General paresis, chr. age, 30; mental age, 11.6.
(*g*) General paresis, chr. age, 41; mental age, 8.9.

GENERAL INTELLIGENCE DETERMINED BY ITS WEAKEST ESSENTIAL ELEMENT

By **Bruce V. Moore**, Research Fellow, Carnegie Institute of Technology, Pittsburgh

In the measurement of intelligence and the interpretation of results, we must have a workable theory, even though tentative, of what it is that we are measuring or testing. We must have a practicable conception of the nature of general intelligence and the processes involved that meet the conditions of our particular tests. Ruml (3) has already discussed the complications arising from inadequate attention being given to our assumptions and hypotheses in mental testing. The following quotations are pertinent: "Relative to the time and number of people devoted to work with mental tests, the results have been astonishingly meager in theoretical value. . . . Extensive collection of data through mental tests began without the necessary antecedent and contemporaneous development of point of view, hammering out of contradictions in concepts and hypotheses, and elimination of ambiguities in common everyday words and ideas. . . . It is probable that many of the failures of mental tests can be traced to our present inadequate theoretical foundations."

Opposing theories have been set forth. Spearman (6) (7) would use those tests for determining general intelligence which have the highest correlation with each other. His theory of a "General Intelligence Factor" or "Intellective Energy" and a hierarchy of special mental abilities with the rank of each depending upon the relative amounts of the general factor in each special ability, was based largely on the high correlation among some mental abilities as shown by tests. Thompson (10) has shown that Spearman's results may not be due to a general factor. According to Spearman's theory, it would be logical to believe that to test a general factor of intelligence entering into special abilities, we should use those tests showing a high correlation due to the presence of this common factor.

Stern (9) has presented a different view as follows: "We must combine together tests that correlate *less* with one another than each one of them correlates with estimated intelligence, and that combination whose amalgamated rank-order shows the highest and most consistent correlation with estimated intelligence is the system of tests that we seek." The theory underlying this statement has been set forth by Thorndike (11). General intelligence is composed of a large group of factors or abilities; and instead of distinct types of intellect and character, there is a gradation around "one-mediocrity." "In our ratings of men we unconsciously strike a sort of average of his abilities in learning, thinking, and acting. The source or cause of this average ability is what we really have in mind when we speak of his intelligence." (12)

Various other investigators have contributed to the discussion and facts concerning general intelligence.[1] All the results and hypotheses might be considered as indices of the nature of general intelligence; but it is difficult to bring them together into a consistent conception of general intelligence that will in turn help us interpret them. We accept Thorndike's theory of a multiplicity of abilities with variations around a single type of 'one-mediocrity' as the most consistent with the facts; and it seems to be coming into general acceptance. Taking this theory as a basis, we shall try to elaborate a working conception that will help us understand the facts, and act as a basis for further study. This conception is presented most easily in diagrammatic form; but first we shall state some of the premises.

---

[1] Wells (14) gives a warning in stating that we may give the same test to two people and yet not test the same thing. "Because Peter is ten per cent better in memory *test* than Paul, does not say that it is due to a ten per cent superiority in the *same* abilities as Paul's." (Italics mine.) He considers general intelligence as mental capacity for adaptation to life, and states that some mental factors most important for this adaptation and success in life are not reached even by Binet tests. "The regularity with which a voluntary task is performed, the attentional control over it, and its freedom from subjective interference are more important to observe than the absolute efficiency in some task."

Burt (2) gave a variety of tests to about a hundred subjects, and in discussing the results he states that the rank order of the tests according to the attention required corresponds fairly closely to the order of correlation with intelligence and with the hierarchy. He states further: "It is one feature or function of attentive consciousness in particular which forms the basis of intelligence—namely, the power of readjustment to relatively novel situations by organising new psycho-physical co-ordinations. . . . High intelligence seems to mean high capacity for continually systematising mental behavior by forming new psycho-physical co-ordinations, older co-ordinations being retained."

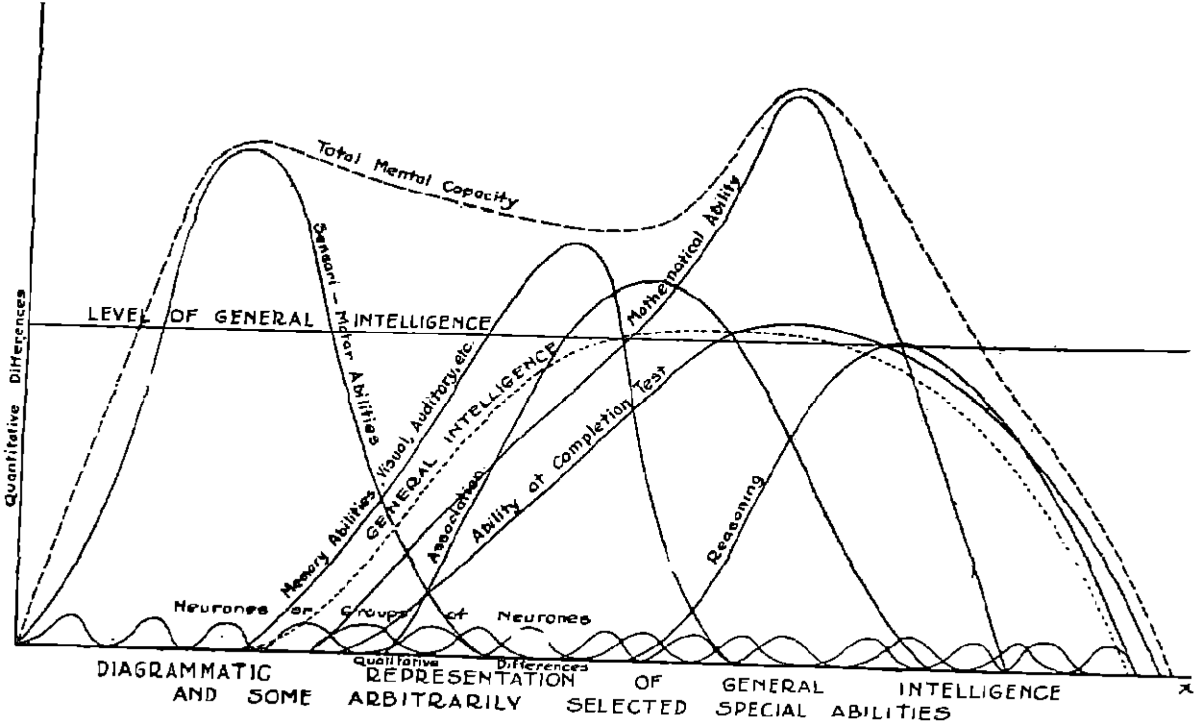
Brown (1) states from his results that correlations may be very low even within a set of mental tests which appear to measure closely related mental abilities. Also, "in homogeneous groups of subjects there is no positive evidence of the existence of one 'central factor' to which the correlations between the individual mental abilities may be regarded as due."

Simpson (5) concludes from his results: (a) "We may well look to language tests in some form to furnish good tests of general intelligence." (b) There is "justification for the common assumption that there is close inter-relation among certain mental abilities, and consequently a something that may be called 'general intelligence': and that on the other hand certain capacities are relatively specialized, and do not necessarily imply other abilities except to a limited extent." (c) There is "no justification for the view that 'general intelligence' is to be explained on the basis of a hierarchy of mental functions, the amount of accuracy in each case being due to the degree of connection with a common factor" (in Spearman's sense). (d) Tests most intimately related to other abilities are selective thinking, memory and association, quickness and accuracy of perception, motor control, and sensory discrimination, each in the order named.

The numberless neurones combined into many situation-response units have still further numberless possible combinations and permutations; but they tend to become grouped and function in the form of what we discriminate as particular abilities. This integrative action of the nervous system is clearly described by Sherrington (4). It was also referred to by Burt, quoted in the note above. Consciousness at any moment, then, usually involves several empirically analyzable functions or abilities, such as memory, association, or rather such as memory of the multiplication table, discrimination of the meaning of the words in a problem in multiplication before us, the association of the problem with processes of multiplication, etc. At different times, different functions or abilities are emphasized or used. Related abilities,—that is, certain related groups of neurones,—are aroused also, and are either used or are ready for use. The state of consciousness might be compared to a correlation surface or a group of mountains with the clearest consciousness represented by the highest peak in the center and the associated foot-hills grouped around it and being smaller the farther they are away from the focus of consciousness. It might be compared better to the surface of a body of water in which the highest wave was the center of consciousness, which was always shifting its position. General intelligence consists of a multiplicity of abilities, or the capability of functioning of several empirically or experimentally analyzable functions or abilities. The many abilities tend to act as a whole, and many are involved in all states of consciousness. For example, if there were no memory, or more strictly speaking, absolutely no retentiveness in any individual, it is difficult to conceive how there could be any intelligence. The intrinsic capacities or potentialities of the various functions or abilities are fixed and unchangeable, although the focus of consciousness playing over them is constantly shifting its emphasis from one function to another.

As we now test general intelligence, we may not, and probably do not test the capacity of the total multiplicity of abilities. We test those that correlate with estimated intelligence, which is general adaptability to environment and ability to get along most satisfactorily in this life as now ordered. We tend to emphasize the importance of the higher processes and test them as representative of intelligence. By higher processes we mean those that have a greater complexity of structure or integration, and have been formed latest in the process of evolution. As these higher abilities are more newly formed and less stable processes, they are somewhat limited in any individual, and are not so common in society. Thus, they are at a premium in the world of action. Their limitations in each particular mind mark the limitations of general intelligence in that individual. The lower order abilities are needed, doubtless, to make up general intelligence, but being developed earlier, they are more stable, their capacities or possibilities are greater, and they seldom limit general intelligence by not being able to contribute their proportionate part necessary for general intelligence. They are not less in capacity or perfection than the higher abilities; they are probably more able to fill their place in general intelligence. But general intelligence can never be greater than the limit set by any necessary factor, ability, or group of neurones when this factor has reached its particular maximum of efficiency in its part of general intelligence. In this respect, *general intelligence is similar to a chain, which is no stronger than its weakest link.* Though there be some compensation or vicarious functioning, it is not necessary for all elements to be weak to weaken general intelligence. However, most functions have a larger capacity or possibility of function than that required by the maximum of general intelligence, since general intelligence is limited by some relatively weak ability. In such cases a measure of general intelligence as it is commonly made does not give a picture of the total mental capacity. Only a psychogram representing the individual capacity of each special ability isolated and measured would give us the total picture. Moreover, by measuring special abilities, we get individual differences, and can thus differentiate individuals; but we are not sure that these differences are differences in general intelligence. It is conceivable that a factor or ability most essential to general intelligence in that it is predominant in quantity, might correlate very little with general intelligence, if it has capacity still greater than its essential contribution to general intelligence requires. Thus a simple factor which logically seems essential to general intelligence, when tested, often shows little correlation with intelligence. This conception may explain some of the unexpected correlations and unexpected lack of correlations in the results of various investigators. Individual differences in the capacities of special functions may hide or make indistinguishable the correlations with general intelligence of special functions tested. The exaggerated development of any particular special capacities may not increase their contribution to general intelligence in the same proportion; that is, it may minimize their importance relative to their exaggerated size, and thus minimize their correlation with general intelligence. The rank of some functions would be lowered and the rank of others raised, so that a "hierarchy of functions" might be or might not be made to appear. There is not a hierarchy of functions, except as they may be grouped; and there is not a common element, except as they are similar in that they are related in some way. There may be a common element in similarity of modifiability or in ability to function, which in cumulation determines general intelligence. General intelligence can be measured not necessarily by its most common function or element, or the one contributing the greater proportion, but by *the one whose maximum possibility is required.* A more complete picture would be, of course, a psychograph of its factors.

This conception is presented diagrammatically on page 160. The extent or length of the x-ordinate or base of the diagram represents the extent of qualitative variety or the multiplicity of abilities of the total capacity of an individual. The height of the diagram of any particular part represents the quantitative aspect or the amount of any particular ability. The small figures or curves on the base line are diagrammatic representations of neurones, or rather groups of neurones which function as the physiological elements of the particular abilities. A few of the particular abilities are diagrammatically represented by the curves similar to frequency curves. General intelligence might be represented by the dotted line touching the tips of the particular abilities entering into general intelligence as it is usually considered. The measure of general intelligence might more consistently be represented by *the elevation of a straight line drawn so as to just touch the top of the curve of the weakest necessary ability.* The broken line represents the total capacity of the mind as it might be found to be by a complete survey for a psychograph. The


Quantitative Differences
LEVEL OF GENERAL INTELLIGENCE
Total Mental Capacity
Sensori - Motor Abilities
Mathematical Ability
Memory Abilities, Visual, Auditory, etc.
GENERAL INTELLIGENCE
Association
Ability at Completion Test
Reasoning
Neurones or Groups of Neurones
Qualitative Differences
x
DIAGRAMMATIC REPRESENTATION OF GENERAL INTELLIGENCE
AND SOME ARBITRARILY SELECTED SPECIAL ABILITIES

high peak in the center would indicate that this particular individual was a mathematical genius. A test of his mathematical ability would differentiate him from other individuals, but would not represent his general intelligence, which was determined by his weakest ability necessary for general intelligence. The completion test would be a good measure of his general intelligence. The diagram might be three dimensional in the form of a correlation surface appearing as a group of mountains and hills of particular abilities.

REFERENCES

1. BROWN, W. Some Experimental Results in Correlation of Mental Abilities. *Brit. J. Psy.*, 1910, vol. 3, pp. 296-322.
2. BURT, CYRIL. Experimental Tests for General Intelligence. *Brit. J. Psy., 1909,* vol. 3, pp. *94-177.*
3. RUML, BEARDSLEY. The Need for an Examination of Certain Hypotheses in Mental Tests. *J. Phil. Psy. and Scientific Method,* Jan. 29, 1920, vol. 17, pp. 57-61.
4. SHERRINGTON, C. S. *The Integrative Action of the Nervous System,* 1906.
5. SIMPSON, B. R. Correlations of Mental Abilities. *Teachers College, Columbia University Contributions to Education* No. 53.
6. SPEARMAN, C. "General Intelligence" Objectively Measured. *Amer. J. Psy.,* 1904, vol. 15, pp. 201-292.
7. —— —— and HART, B. General Intelligence, Its Existence and Nature. *Brit. J. Psy., 1912, vol.* 5, pp. 51-84.
8. —— ——. Some Comments on Mr. Thompson's Paper. *Brit. J. Psy.,* 1916, vol. 8, pp. 282-284.
9. STERN, L. W. *The Psychological Methods of Testing Intelligence.* Warwick and York, 1914.
10. THOMPSON, G. H. A Hierarchy without a General Factor. *Brit. J. Psy.,* 1916, vol. 8, pp. 271-281.
11. THORNDIKE, E. L. *Educational Psychology,* vol. III. Teachers College, 1913.
12. —— ——. Intelligence and Its Uses. *Harper's Magazine,* Jan., 1920, pp. 227-235.
13. —— ——. Significance of the Binet-Simon Tests. *Psychol. Clinic,* 1916, vol. 10, No. 5.
14. WELLS, I. Principle of Mental Tests. *Science,* 1913. n. s. vol. 38, pp. 221-224.

EFFECTS OF VARIOUS MODES OF PUBLIC READING[1]

By Charles H. Woolbert, Ph. D.

This study is an attempt to bring the numerous problems of oral expression and public reading into the laboratory. For the beginnings of such an experiment the most promising place is the laboratory of psychology inasmuch as speech is intended to reach the minds of auditors who are expected to react to the stimulation furnished by the voice of the speaker or reader. Thus the problem deals with auditory stimulations and with behavior activities of listeners. The task to be met is that of finding a method whereby the stimuli can be controlled and receptivity measured; that is, of insuring an adequate description of the effect of the voice of the speaker or reader upon the responses of the auditors.

The method adopted after many preliminaries was based upon a study of the relation between *changes* in the use of the voice and specified responses of auditors; that is, the effects of various modes of public reading, employing *different combinations of changes* in the use of the attributes of sound while reading. These different combinations offer various *modes of reading*. The attributes of sound are pitch, intensity, time, and timbre (sometimes called quality when applied to the voice). While all speech that is meaningful to auditors employs constant changes in the use of these four attributes, still there are no results obtained under laboratory conditions that show which are most significant or what relation they bear to each other. A restricted series of modes of reading was selected for use throughout the study, limited to certain *gross degrees of change* in pitch, time, intensity, and timbre. A description of these modes together with the method of evaluating results from their use, is given later.

Changes in pitch, time, intensity, and timbre can be studied as general and total or as finely modulated and of narrow degree. Again they can be studied as affecting individual words, sentences, paragraphs, or larger units of discourse.

---

[1] From the Psychological Laboratory, Harvard University.

The type of change studied here applies to *large units* of discourse, in this case passages of twenty minutes in length. No attention is here paid to the modulations occurring between words, or sentences, or units of small measure. A certain "tone of voice" is selected and used throughout a twenty-minute passage.

THE ATTRIBUTES OF SOUND AS ELEMENTS OF CHANGE IN READING

Pitch Changes. When employing changes in Pitch as the basis of investigation, it is necessary to hold in mind that these are of at least three different kinds: (1) the slide within the individual syllable, (2) the step from one syllable to the next, and (3) the general pitch level on which a passage or a composition is uttered. The first two are continual changes required in normal speech; the last change is only occasional. When change of pitch is referred to in this paper, change of the slide within the syllable and from one syllable to the next is intended.

Time Changes. Changes in time are chiefly in the length of the vowel parts of words, i. e., the tonal part as against noises. "Gauthiot's[2] records show that there are two main classes of vowels in respect to duration, long and short, whose average lengths bear the remarkably constant relation of 4:2."

Changes in Pause. A cause of the frequency of duration change is that pauses must be included. Pauses—periods of silence—are necessary because of (1) the need of a break between syllables and words, and (2) because of the demands of breathing. Accordingly ,in speaking here of changes in time we always include changes in pause; pauses between syllables, words, phrases, clauses, sentences, and the major divisions of composition.

Intensity Changes. Changes in tone intensity are also very common owing to the fact that accentuation is not possible, in English, without a change in intensity on the accented syllable; for accent, in this language at least, is largely a matter of a rise in pitch and an increased energy of utterance. Also emphasis, as distinguished from accent, is almost always accompanied by an increase or decrease in intensity.

Qualitative Changes. Changes in quality (timbre) are least common of all. Changes in quality are probably more influential than any other type of change in revealing the emotional state of the speaker. Emotional states involve muscular and vascular changes in the throat and pharynx, and thus the body of overtones is greatly affected in number and intensity, producing attention changes in quality.[3]

---

[2] Gauthiot, De l'accent et de la quantite en literanien, *La Parole*, 1900, II., 143. (Cited by Scripture.)

Overlapping Effect of the Four Attributes. In every manner of reading that involves a maximal change of one of the elements, there is a certain amount of overlapping. When the extremes of change in pitch are used, there is a necessary change in quality. Very high notes, lacking certain overtones, produce a different quality impression from tones possessing these overtones. A prolonged tone, especially a tone of speech necessarily involving a slide, tends to carry the slide on during the prolongation, thus involving a change in quality. A tone of strong intensity is kept loud most easily by being also high in pitch. Tones of "chest resonance" are prevailingly lower in pitch than those of "head" quality; at least the majority of people find it easier to form them in these combinations.

By isolating one attribute at a time, giving it an extreme degree of change while holding the others at a median degree, we get 4 modes of reading; by doing the same with no change of each of the attributes we get 4 more modes; by using a median change of all 4 we get another; by using an extreme degree of all four we add one more; and, finally, by employing no change at all in any of the attributes we add the last; making in all 11 modes of reading. These represent only gross situations, and so are capable of control at the hands of a trained reader. Following is a specific designation of each of these modes:

Mode 1: Median change of all four attributes.
Mode 2: Extreme change of Pitch; Median change of other three.
Mode 3: No change of Pitch; Median change of other three.
Mode 4: No change of Time; Median change of other three.
Mode 5: Extreme change of Time; Median change of other three.
Mode 6: No change of Intensity: Median change of other three.
Mode 7: Extreme change of Intensity; Median change of other three.
Mode 8: No change of Quality; Median change of other three.

---

[3] Blanton, S. *Quart. Jour. of Public Speaking,* Vol. II. "The Voice and the Emotions."

Mode 9: Extreme change of Quality; Median change of other three.

Mode 10: Extreme change of All Four Attributes.

Mode 11: No change in any of Four Attributes.

MODES OF READING

Mode 1. Median Change of All Attributes. In this mode all changes are of a median degree, but are constant in all four of the attributes of sound. Extremes of all kinds are avoided. The general effect of this type of reading is that produced by a congregation reading a Psalm in unison. There is a lack of what might be called "brightness," "animation," "alertness." In general it would be called rather dull reading, yet far from unintelligible and not at all bizarre or unusual.

Mode 2. Extreme Change of Pitch.[4] In this mode the voice travels to extremes of height and depth continuously, in the use of the slide on the syllable and in the step from syllable to syllable. Changes are extreme not only in distance on the scale of pitch, but are frequent in occurrence, no passage of more than three or four syllables being without its abrupt and wide change. This mode is not altogether removed from meaningful and intelligible speech; for not a few sentences resemble a type of speaking frequently heard in public address and in private conversation. Yet the continual shifting of the voice up and down to extreme depths and heights produces an effect of monotony that is marked.

Mode 3. No Change of Pitch. In this mode the reader eliminates all inflection of the voice upwards or downwards on the scale whether in slide, step, or general level. In the first place, the level is approximately at the middle height of the reader's range. Then, all slides also were eliminated, and likewise the steps; so that every sound was of the same level as every other sound. This type of reading is far from meaningless and is by no means strange to listeners who are accustomed to the "intoning" customs used by certain ecclesiastical sects.

Mode 4. No Change of Time. In mode 4 every syllable was rendered with the same duration as every other syllable. Moreover, the pauses were of the same length, exception being made when breath had to be taken. A manner of speech employed by certain foreigners in learning the English language approximates this; also, some speakers in their anxiety to be distinct use this mode. A most noticeable effect is the staccato involved. This mode, though, is not without meaningful emphasis, inasmuch as the emphasis of slide and intensity are retained to a median degree.

---

[4] A characterization like this assumes, it is to be remembered, a median degree of change of the other four attributes.

Mode 5. Extreme Change of Time. Widely varying degrees of change in time were applied not only to the vowel sounds, but to the pauses. Still the prolongations in this mode were not so great as is sometimes heard on the stage and in impassioned speaking, where with such an extreme manner must go a wide range of pitch or an amplification of the quality or an increase in the intensity that would constitute an error in the use of the median range of these three demanded by the premises. This is a relatively common mode; "animated" conversation is almost always markedly broken in time.

Mode 6. No Change of Intensity. Reading on one level of intensity encounters a difficulty, in that there is need of eliminating all intensity factors from accentuation; and that is not easy. When all syllables must be made equally loud, accent disappears. Thus, choice had to be made between making unaccented sounds louder or accented sounds quieter. The latter alternative was chosen, for the reason that otherwise the whole effect would be merely that of blatant shouting. As a consequence, the reading in this mode gave the impression that the reader was very tired or uninterested in his task. Also when such accentuation by intensity change is eliminated, the listeners get the impression that the words are not being pronounced correctly and so reveal distress at first sound of the mode. Later they seem to be relieved.

Mode 7. Extreme Change of Intensity. The simple formula for describing this mode of reading is to state that the sounds that in "normal" speech are the loudest are here made very loud; those that are of a median degree of loudness ordinarily are made loud; the sounds of least intensity in every-day speech remain about the same, for they are merely audible then and cannot suffer lowering of intensity without disappearance. Listeners seemed to be benumbed by this manner.

Mode 8. No Change of Quality. To accomplish this, one system of resonance, or complexity of sound, had to be decided upon arbitrarily from the several possible in the use of a trained voice. The quality chosen is that known as "oral" in terms of elocution. It is a quality that derives its resonance chiefly from the bones high in the head. It is rich in high overtones, but poor in the lower ones that add much to impressiveness in speech. All resonance from the "circle of the pharynx" is reduced to a minimum, if not eliminated altogether. The effect of this manner is to make the reader appear sick or weak.

Mode 9. Extreme Change of Quality. In this mode various complexities of tone are produced by "placing" the voice so that there is a continual change in the various resonance-producing chambers of the head, throat, and chest. The effect is what is loosely called a continuing change of "emotional" tone. This makes it a rather common manner of speaking. It carries the virtue of "interestingness."

Mode 10. Extreme Change of All Four Attributes. Here we have the mode that makes variation its chief quality. If variation may be predicated *a priori* as a value in reading, then this ought to prove an impressive and an interesting manner. Every kind of change is employed and in profusion. No one kind of change or lack of change was employed for any length of time. While great variety of change prevailed, yet only those changes were chosen which the reader thought were in conformity with the social speech habits of the type of auditors who served for subjects. It is a mode that empirically would be judged effective in making commonplace statements interesting and easy to listen to.

Mode 11. No Change in Any of the Four Attributes. This manner involved complete uniformity of pitch, time, intensity, and quality—the extreme of monotony. It inhibits the apprehension of the logical meanings of word-sounds. When all syllables are given the same duration and the same intensity, in addition to monotony of pitch and absence of change of quality, the listeners seem to have difficulty in catching the words. Yet this applies at the beginning of the reading more than later.

THE READING

The reader made no attempt to add to the effectiveness of his reading by any other means than the use of the voice. No gestures, postures, or facial expressions were intended; all such devices for inducing responses were scrupulously and successfully, it is fair to say, eliminated. Sitting comfortably in his chair, most of the time leaning back, though occasionally leaning forward on the desk for relaxation, he read in the manner he had prescribed for the day's reading.

At the close of the reading the reader shut the book, and announced that that was all for that reading. Then an interval of a few minutes was allowed, varying from two minutes to five. During this interim the experimenter wrote on the blackboard the title of the second story to be read during the next half of the period, together with the names of the characters that appeared in that story. The auditors were free to move about the room or to remain in their seats. The only prohibition put upon them was that they must not talk about the story just read or about the characters in it or about stories of which it reminded them. In fact they were prevented from making any reference or allusion to it that might establish or strengthen associations.

At the close of the second reading, the experimenter passed out the papers on which the reports were to be written. These papers were folded, numbered on the back, and dated with the date on which they were to be filled in. Subjects were instructed not to unfold them, and under no circumstances to look at the contents until ready to write out their report. The papers were presented in such a way that there was no occasion for looking at their contents. There is every reason to believe that this instruction was carried out to the letter. When the papers had been given out, all the auditors left the room.

Length of Reading. The reader stopped arbitrarily at the close of the twenty minutes. With two readings a week for each subject and with two readings a meeting, a large number of stories had a place in the experiment.

The reports were returned to the experimenter any time after the fifth day.

REPORTS

Purpose. The purpose of the reports is to provide a measure of impressiveness of the different modes of reading, and to give this measure in mathematical terms. To do this required a new method with very little precedent as guide.

Tentative Method of Completion Test. In the first test to measure retentiveness of impression according to modes of reading, the subject was merely asked to write down what he remembered. No restrictions were placed upon the amount to be written, whether a brief summary or an extended narrative. As a consequence the reports differed so widely in the number of words employed and in minuteness of detail that they were not at all comparable.

Method Revised Further. To remedy this, a fixed schedule of instructions was given the subjects, as follows:

1. Identify and characterize the different personages.
2. Describe briefly the places pictured.
3. Give a consecutive account of the events of the narrative.
4. Outline the conversations given.

Again the differences in the fulness of the reports of subjects were too great to make the results at all comparable; some subjects wrote volubly on these points, others in the most laconic manner possible. So this device was abandoned.

Method Finally Chosen. To remedy this defect, the system of "leads" was adopted. The "lead" is a partial or complete declarative sentence so selected that it brings to the observer some situation incident, or fact concerning the story. When the sentence is incomplete, then there is placed in parentheses some instruction as to what to do to complete the sentence. If the sentence is complete, however, then in the parentheses is given some instruction as to the offering of details, consequences, or characterization. Following is a sample of the type of "leads" used in the experiment:

THE INN

The inn of which we speak possesses many attractions (name 6).
The present proprietor is Naum (describe him).
The previous proprietor had been Akim (describe him).
Akim kept the inn (how?).
The land belongs to Elizaveta Kuntz (characterize her).
Relate the affair between Akim and Dunyasha, the servant.

Another example of this type of "lead":

FATHERS AND CHILDREN

The governor in the narrative was to be investigated (for what?).
Matyei, sent to investigate, was (characterize).
Matyei stunned his subordinates in the characteristic Russian way (how?).
Sitnikoff the dilettanti took them to the home of Madame Kukshin (describe her).
The present domestic status of Madame K. was (what?).
The three men were induced to go by (what promise?).
Madame Kukshin talked of (detail her interests, using proper names).

LIEUTENANT ERGUNOFF

The Lieutenant was treated by the ladies from Riga (how?), (give specific details).
One day the Lieutenant, wearied by (what?), (did what and with what results?).
The next day the Lieutenant received a letter (give contents).
Upon his arrival next day (describe his welcome).
(Then what happened? Relate in detail.)
(Give a visual and an auditory picture of the "Humming-Bird.")

LEADS

Some Problems Involved. In devising leads several considerations have to be kept in mind, in order to make the data comparable:

(1) Nothing subjective was asked for beyond interpretations as to character and the meaning of events in the story. In the preliminary stages of the use of this method occasional descriptions and characterizations were demanded; but results demonstrated that individual differences in responding to such leads were so great that the results of the various subjects could not be compared.

(2) The matter asked for in the leads was uniformly concrete.

(3) Leading questions, those that can be answered with Yes, or No, were entirely eliminated. Also questions were so devised as to make difficult the inferring of the answer from the nature of the question.

(4) The leads were chosen in such a way that the total effect of them was to produce a running summary of the story in its chief details.

Number of Leads. According to this method there is no fixed number of leads for a report. In fact, the limiting of the number of leads to some specified figure would bring an error into the records, for the reason that there are differences in style from one story to another of enough significance to make necessary a flexible joint in the method. Thus the choosing of the number of leads is entirely a matter determined by the nature of each story or passage. Especially determinative of the flexibility is the matter of concreteness in the passage. In particular, events in time make serviceable material for leads. The number of leads presented on reports differs from 8 to 35.

SCORING

Determining the Value of a Point. In general every recorded perception of a process was counted a point. Then when characterizations were demanded, if the number of these was specified, the fixing of the value of a point was a function of the number demanded; if the number was not called for, the experimenter had to decide what kind of answer constituted a point. In general a "point" is to be defined here as a recorded fact of an event, a characterization, or a description. As nothing but objective matter was asked for, the factual basis of points made scoring reports reasonably accurate and uniform.

Specific Answer. The answer to the lead must satisfy the specific question involved before any credit at all is given.

No Overlapping. The number of points given any specified lead must be only such as are needed to tell the facts called for without *overlapping.*

Essential Details. Only such details as are essential to the event, characterization, or description are accepted.

Verbosity vs. Conciseness. Neither verbosity nor conciseness are determining factors in scoring points. Some reporters are habitually wordy, others are by habit laconic. Regard must be had for this difference in estimating the number of points to be scored. A single word can carry the evidence that the reporter has retained the essential fact while sentences may be employed by another reporter without adding to the evidence for retention. This was kept in mind consistently in making out the scores.

Lack of Specificity in Answers. No credit is given for statements that do not specifically answer the question asked.

Guesses. Subjects are warned against offering guesses; they are encouraged to leave blanks rather than make replies the correctness of which they cannot substantiate. The important consideration in the understanding of their task is that they are not to make guessing a practice for the sake of shielding themselves from the effort necessary to answer with assurance, or for the sake of a full report.

The Total Score. The total score for a report is a fraction which uses the whole number of points assigned to that report for a denominator and the total number of points scored for a numerator. For purposes of uniformity this fraction is treated in the results on a percentage basis.

Mechanism of Scoring. As a measure for stabilizing the judgment of the scorer, all scoring was done with the reports of one selected painstaking subject as guide.

The first step in scoring is to estimate each lead separately, to determine how many points are needed to represent full retention for this lead. The number thus secured is then placed beside the lead at the left side of the paper. After all of the leads have been thus estimated, the total is found and set down at the top of the sheet for the denominator. The score of the paper, however, is not yet settled finally; two or three more papers are graded on this basis before arriving at a final judgment as to the weight of each lead and so of the size of the denominator. This reading of additional reports sometimes shows errors in the preliminary judgment, or reveals that the replies in the report of the guide auditor are not precisely representative. Revision is then made of those thus far graded, and with the revised denominator as a final judgment the scores are made out for all the reports and entered at the top of each sheet both as a fraction and as a decimal.

A TYPICAL REPORT GRADED

Reading A2

Wide Range of Pitch

MUMU

(Vol. 2, to p. 218)

(2) IN MOSCOW LIVED WITH her retainers and
2 a Russian mistress (widow) servants

(5) AMONG THEM WAS GERASIM WHOSE DUTIES WERE (SPECIFY FULLY)
3 to sweep out the courtyard/fetch water/and generally guard the premises/

(3) GERASIM HAD PROVED HIS COMPETENCE BY (SPECIFY THREE DEEDS)
3 not letting any dirt appear in the yard/if the cart for fetching water stuck in the mud he extricated it by his own strength/knocking the heads of two hapless robbers together/

(4) IN CONVERSATION WITH GAVRILA THE MISTRESS DECIDED TO
3 marry off one of her dependents, a drunkard/in the hope of reforming him/and settle on Tatyana as his destined wife/

(3) GERASIM'S COURTSHIP OF TATYANA HAD CONSISTED OF (3 ITEMS)
2 presenting her with little tokens as she passed him in the courtyard/not permitting her to be made sport of in his presence/

(2) THE MISTRESS UPON HEARING OF GERASIM'S COURTSHIP WAS
0 I did not think the Mistress knew anything of Gerasim's courtship.

(3) KAPITON'S REPLY TO GAVRILA CONCERNING TATYANA WAS

3 that he would undoubtedly be killed by Gerasim/but that he was always doomed to ill luck/and might as well marry. He liked Tatyana herself/

(2) THE EFFECT OF THE COURTSHIP UPON TATYANA WAS

2 merely to frighten her to death/and to acquiesce in the courtship for that reason/

Score—18/24; .75

Significance of Symbols. The capital letters used as designations of the readings represent the mode in which the reading is done. (pp. 126-7.) The Arabic numerals used in connection with the capital letter represent readings 1 and 2, the first and the second respectively during the hour of meeting. The text in capital letters represents the leads given out on the reports; the rest of the text is the answers given by the subjects. The Arabic numerals in parentheses at the left represent the total number of points allowed to the lead beside which they are placed; the numerals not in parentheses are the score given to the answer. The fraction at the bottom of the report represents the total score, which is also given as a decimal. The use of the bar (/) in the answers marks the division into points; they are inserted to indicate the basis of marking off into points.

Preliminary Discussion

Constant Factors. The attempt was made in this experiment to keep the conditions as similar as possible to the conditions which prevail in places of public meeting. The typical "audience" was of three, occasionally a fourth being present, and at rare intervals five. Audiences of two were very few in the total number of tests.

Thus the social environment of an audience was achieved in all cases. The tests were held in Room 27, Emerson Hall, Harvard University. Except when the hour was the first in the morning and subjects showed a tendency to straggle in as when going to public places, they gathered in the corridor outside the room; they were instructed not to enter until the whole group had gathered, or to take their seats and comport themselves in the manner usual to them in going into a place of public meeting. Then as soon as all were in their places, the experimenter took charge of the meeting and directed the activities of the subjects from the platform.

Instructions. The experimenter then read the following instructions:

1. "Attend as you are in the habit of attending at public gatherings.
2. "At the close of the hours you will pass out without being called upon for a report or discussion of the content of the reading.
3. " On the fifth day after the reading, and not before or later, you are to fill out a blank prepared by the experimenter, which will be placed in your hands at the close of the hour.
4. " No feats of memory are sought, and no penalty or reward attaches to the amount you remember; merely set down at the time of reporting what you have retained.
5. "Keep the conditions of reporting as uniform as possible; use the same time of day, the same working surroundings, the same place, the same amount of time, the same degree of care, and the same degree of effort to recall."

During the early tests these instructions were repeated at the beginning of each period; later, paraphrases were given carrying the same implications; while still later nothing more was said as to instructions other than to remind the company that the instructions were the same as before. The subjects were repeatedly instructed that they were to listen as they are in the habit of listening when out in a place of public meeting, and that they were not called upon for feats of memory; they were *merely to listen as they do usually*, and to return an honest report written under circumstances uniformly the same.

The Choice of Reading Matter. The choice of reading matter for this kind of experiment offers several nice problems. In the first place, the matter must be unfamiliar to the auditors; secondly, the reading matter must be of a type to hold interest uniformly keen among the whole group of subjects—narration, and narration that is stimulating, with definite type of characters, and written by a writer conceded widely to be a master of style. The higher the degree of uniformity in certain essential particulars, the more valuable is the matter for the present purpose. Uniformity of rhetorical style, strength of characterization, plot structure, descriptive power, and structural strength, are kept at a maximum by adherence to the works of the same writer. Then, again, uniformity of types of people presented in the stories is an aid to general uniformity of type of matter; also similarity of situation, geographical surroundings, the manner of living of people, and national setting.

With all these requirements in mind, the works chosen for this experiment were the novels and stories of the Russian writer, Turgenieff.[5] Inquiry revealed that the auditors assigned for this experiment had read practically none of his writings. From the first tests it became manifest that the element of interest was adequately provided, judging by the attitudes of the listeners during the reading and by their comment afterwards. Some had to be specifically instructed not to read the author's works during the course of the experiment; they had asked permission to read some of his stories by themselves.

Proper Names. The matter of Russian names was handled in the following manner. At the beginning of the meeting, as soon as the audiences had taken their places and the reader had taken the chair, the title of the story was placed on the blackboard before the audience, together with the names of the characters who participated in the story. This was done with a view to reducing to a minimum any confusion over names so unfamiliar to an American audience as are the names of the Russian people. In no way throughout the experiment was any attempt made to require subjects to remember these names; while every effort was exerted to make recognition of them easy. By putting the names upon the blackboard a visual perception was added to an auditory. Always in the reports these names were so presented by the experimenter that the auditor never had to recall the name in order to indicate which of the people of the story he was discussing or reporting on. How this was handled will appear in the transcript of forms of report presented above.

Resumé of Previous Readings. After the names of the characters of the story had been written upon the blackboard, the reader announced, "I am now going to read you part of a story by Turgenieff, entitled.................." When the story was to be used for the first time, nothing more than this was said. However, when part of the story had been read to a given group, the reader preceded the test with a brief resumé of the story up to the point at which he proposed to begin. This was done to enable every subject to start out with as full an equipment as possible of the story up to that point, so that those of superior retentiveness would not start with an advantage over those who happened to have retained little from the reading preceding. In this way a source of error was eliminated.

---

[5] The edition used is the translation by Isabel F. Hapgood, and published in 7 volumes by Charles Scribner & Sons, New York, 1915.

TABLE I: RECORDS OF IMPRESSIVENESS—SERIES I

20-minute Readings; 2 Readings per Session
Reported after 5-days interval
Auditors Unadapted to Experiment

Subjects

Modes of Reading (See pp. 15–21)	1		11		15		21		22		23		24		25		26		27		
	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	Av.
1	.80	.70	.24	.04	.28	.04	.00	.08	.32	1.00	.08	.19	.32	.70	.60	.81	.92	.52	.60	.63	.45
2	.89	.65	.48	.17	.00	.30	.48	.26	.70	.70	.56	.35	.56	.61	.85	.78	.85	.65	.89	.91	.58
3	.94	.80	.00	.31	.00	.19	.13	.31	.75	.65	.00	.19	.29	.50	.54	.15	.58	.59	.63	.39	.39
4	.94	1.00	.11	.09	.72	.64	.33	.32	.95	.96	.06	.23	.83	.86	.61	.86	.44	.96	.56	.68	.60
5	.83	.88	.60	.00	.88	.83	.13	.00	.93	.88	.13	.53	.57	.82	.57	.88	.75	.82	.78	.41	.61
6	.81	.93	.00	.14	.81	.64	.00	.14	.75	.93	.19	.29	.56	.43	.75	.79	.88	.64	.81	.71	.54
7	.92	.87	.00	.57	.00	.57	.00	.00	1.00	.48	.18	.35	.00	.43	.55	.91	1.00	.87	.00	.57	.47
8	.64	.63	.00	.13	.00	.06	.00	.00	1.00	.83	.00	.19	.91	1.00	.81	.94	.73	.56	.00	.19	.43
9	1.00	.88	.00	.00	.64	.00	.45	.29	.82	.82	.00	.06	.64	.71	.73	.65	.45	.71	.64	.35	.70
10	1.00	1.00	.38	.00	.69	.62	.38	.38	1.00	1.00	.38	.00	1.00	.83	.88	.92	.88	.83	.88	.92	.70
11	.63	.60	.08	.00	.15	.00	.15	.00	.21	.87	.05	.00	.00	.60	.42	.40	.16	.67	.16	.47	.28
Av.	.83		.75		.37		.17		.79		.18		.55		.69		.66		.55		

GRAPH 1:—FROM TABLE I—SERIES I—COMPARATIVE IMPRESSIVENESS OF VARIOUS MODES OF READING

TWENTY MINUTE PASSAGES; REPORTED AFTER FIVE DAYS

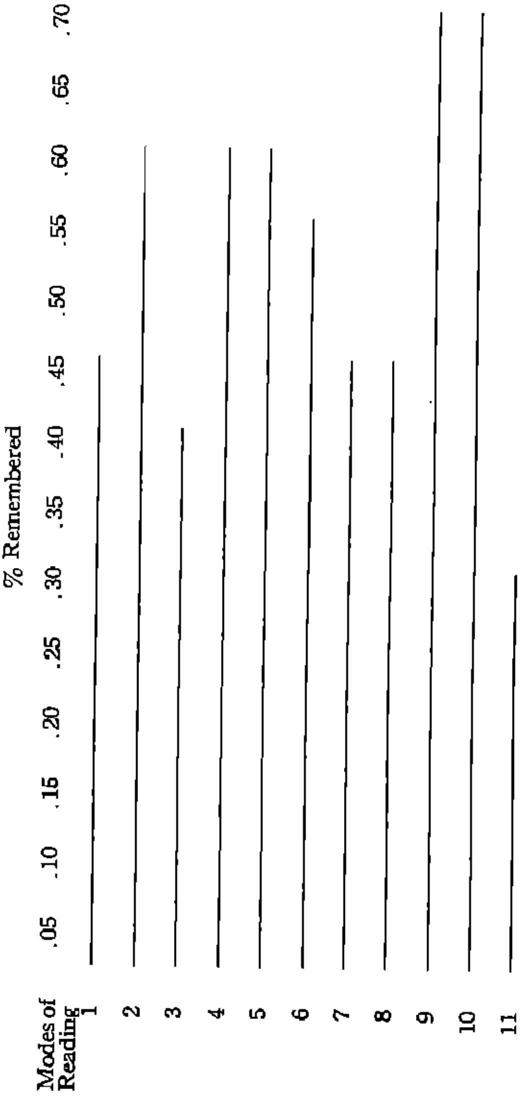
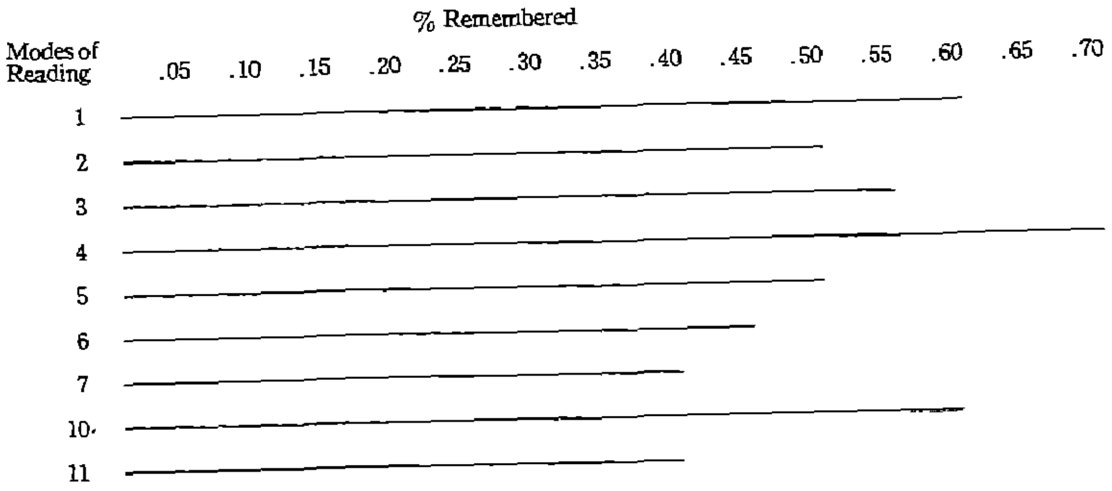


TABLE II: RECORDS OF IMPRESSIVENESS—SERIES II

TWENTY MINUTE READINGS—TWO READINGS PER SESSION REPORTED AFTER FIVE DAYS INTERVAL. AUDITORS ADAPTED TO EXPERIMENT

Subjects

Modes of Reading (See pp. 15–21)	1		11		15		21		22		23		24		25		26		27		
	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	Av.
1	1.00	1.00	.12	.29	.29	.76	.11	.48	.76	.52	.00	.38	.82	.86	.76	.76	.59	.86	.59	.76	.61
2	1.00	.88	.00	.00	.00	.88	.67	.00	.62	.63	.00	.00	.50	.94	.33	.69	.67	.79	.25	.69	.49
3	1.00	.95	.78	.60	.61	.65	.30	.00	.65	.70	.00	.40	.70	.80	.74	.65	.91	.65	.52	.17	.58
4	.84	.87	.82	.70	.83	.87	.23	.00	.63	.70	.20	.57	.82	.74	.97	.78	.60	.87	.83	.83	.69
5	1.00	.78	.35	.56	.29	.61	.12	.00	.79	.67	.12	.00	.65	.50	.76	.61	.65	.56	.65	.39	.48
6	.80	.50	.10	.21	.60	.50	.00	.00	.70	.38	.00	.04	.60	.46	.85	.63	.65	.58	.65	.71	.45
7	1.00	.94	.10	.17	.40	.28	.00	.39	.60	.89	.00	.11	.35	.50	.60	.50	.65	.83	.35	.17	.43
10	1.00	1.00	.42	.00	.31	.86	.00	.00	.58	1.00	.15	.14	.81	.62	1.00	.76	.95	.95	.85	.86	.61
11	.65	.70	.29	.35	.71	.15	.00	.00	.65	.50	.00	.00	.76	.35	.24	.35	.41	.70	.65	.60	.41
Av.	.82		.28		.53		.13		.66		.13		.63		.68		.72		.58		

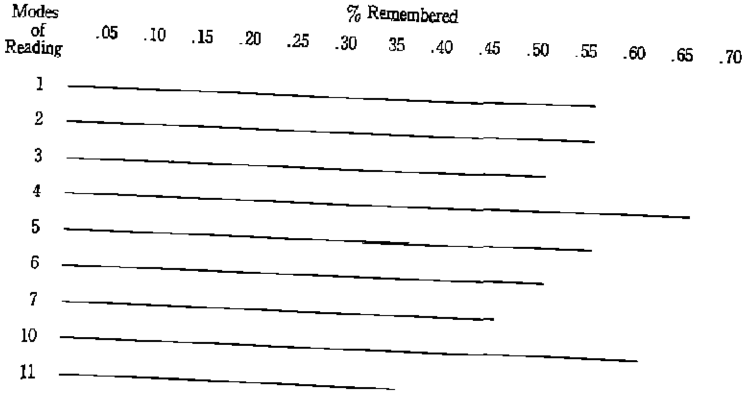


GRAPH 2—FROM TABLE II—SERIES II—COMPARATIVE IMPRESSIVENESS OF VARIOUS MODES OF READING

TWENTY MINUTE PASSAGES. REPORTED AFTER FIVE DAYS

TABLE III—TABLES I AND II COMBINED

RELATIVE IMPRESSIVENESS AS SHOWN BY TWO SERIES COMBINED—FOUR READINGS FOR EACH MODE



GRAPH 4—FROM TABLES I AND II COMBINED

INDIVIDUAL DIFFERENCES IN IMPRESSION FROM NINE MODES OF READING

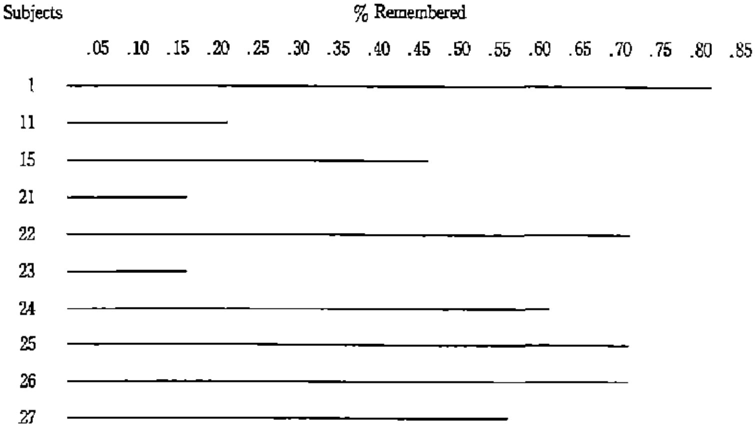
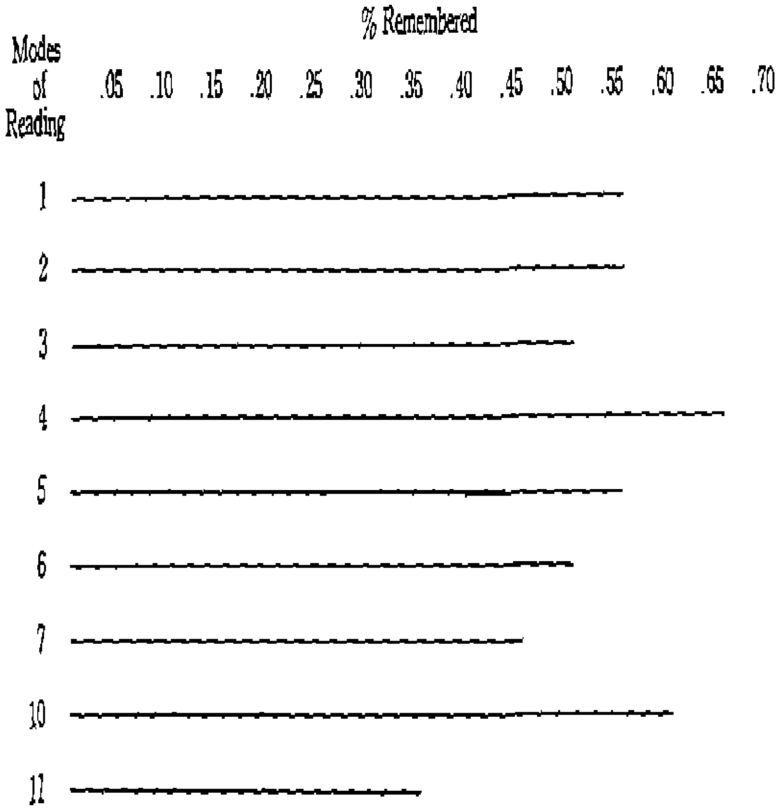


TABLE III—TABLES I AND II COMBINED

RELATIVE IMPRESSIVENESS AS SHOWN BY TWO SERIES COMBINED—FOUR READINGS FOR EACH MODE



GRAPH 4—FROM TABLES I AND II COMBINED

INDIVIDUAL DIFFERENCES IN IMPRESSION FROM NINE MODES OF READING

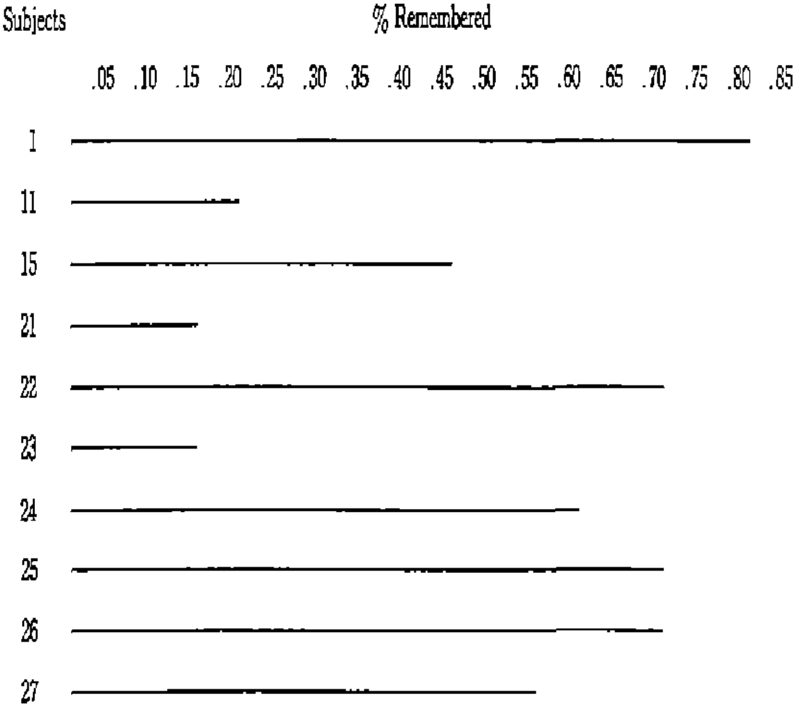
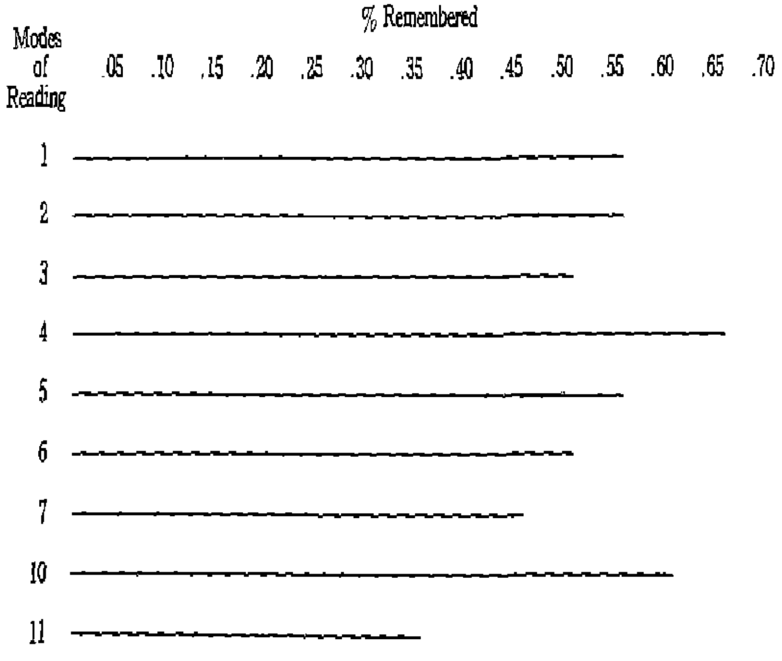


TABLE III—TABLES I AND II COMBINED

RELATIVE IMPRESSIVENESS AS SHOWN BY TWO SERIES COMBINED—FOUR READINGS FOR EACH MODE



GRAPH 4—FROM TABLES I AND II COMBINED

INDIVIDUAL DIFFERENCES IN IMPRESSION FROM NINE MODES OF READING

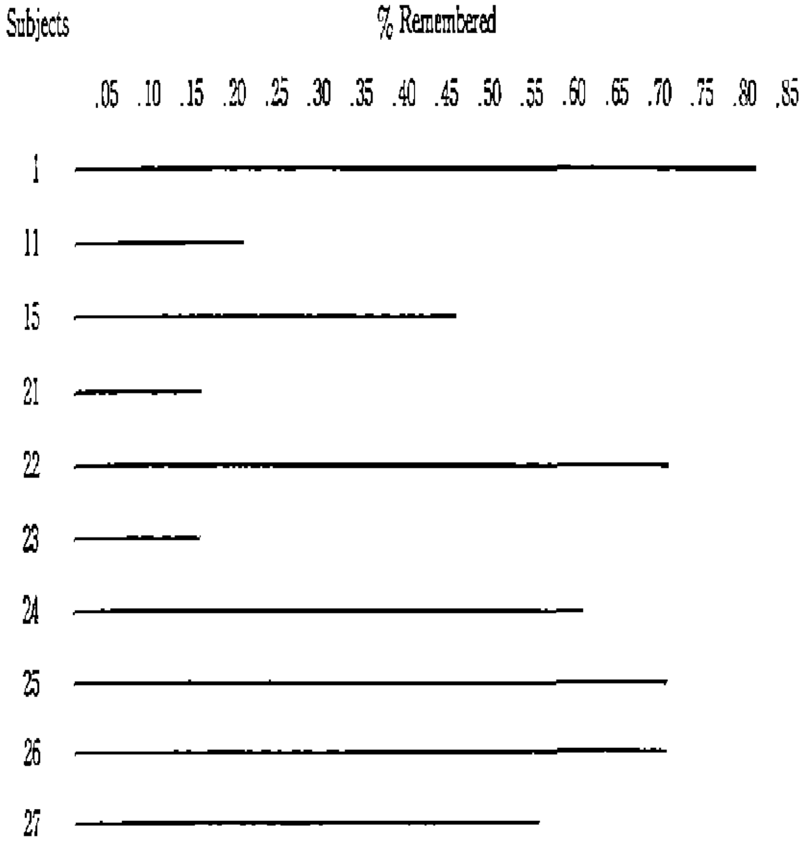
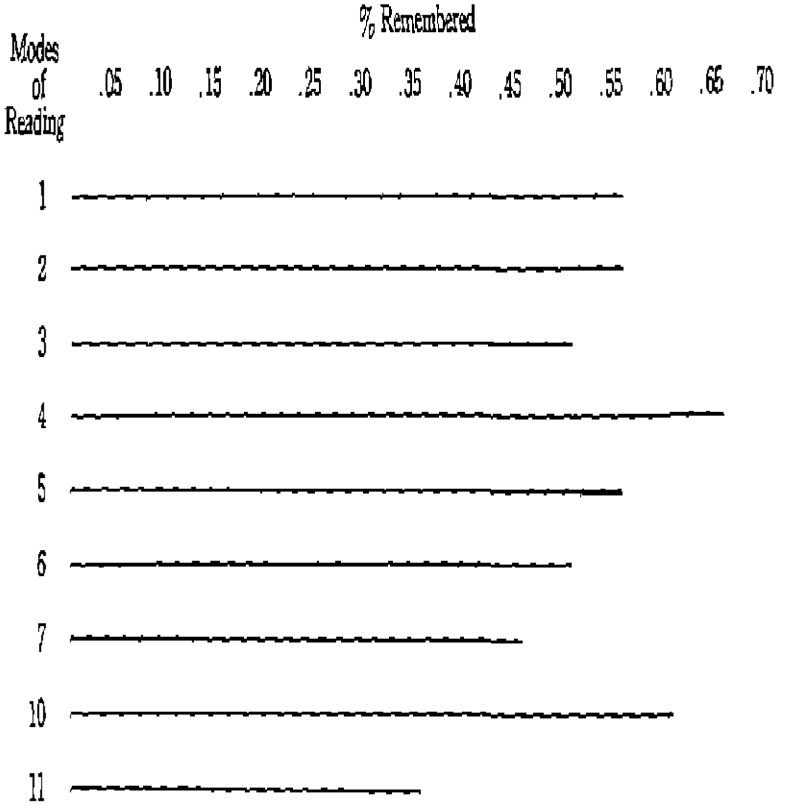


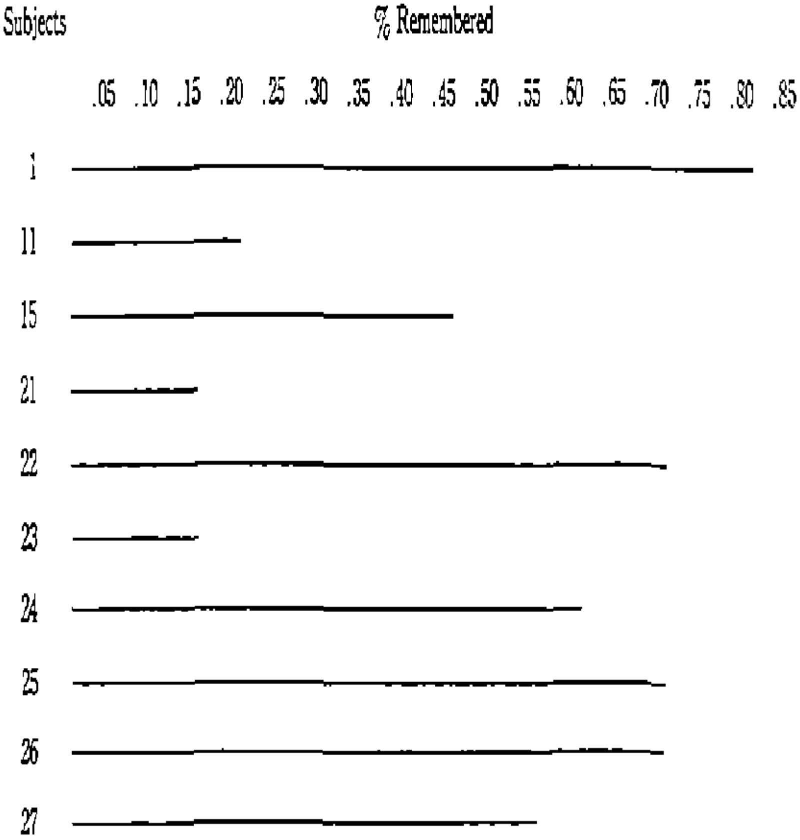
TABLE III—TABLES I AND II COMBINED

Relative Impressiveness as Shown by Two Series Combined—Four Readings for Each Mode



GRAPH 4—FROM TABLES I AND II COMBINED

INDIVIDUAL DIFFERENCES IN IMPRESSION FROM NINE MODES OF READING



DISCUSSION OF RESULTS

The foregoing results will be discussed under the following two heads:

I. The Function of Changes in the Use of the Four Attributes of Sound in Oral Reading.

II. Individual Differences in the Responses of Auditors.

Series I and Series II are alike in technique and material; Series II was added in the conviction that this type of experiment needs a large number of cases for validity of results.

I. FUNCTIONS OF VARIATION IN THE USE OF THE FOUR ATTRIBUTES OF SOUND IN ORAL READING

A. *Function of Median Range of Variation*

a) Mode 1, i. e., Median Range of All Four Attributes, holds the median position in relative impressiveness of the modes of reading employed, in Series II and in the total summary for both series. The same is practically the case for the 11 modes of Series I.

B. *Function of Extreme Range of Variation*

a) Extreme variation of Quality, in the only instance in which it was used, Series I, seems to rank high in impressiveness. Extreme Range of Time (Mode 5) gives results that are inconclusive; while Extreme Range of Intensity (Mode 7) is low enough to suggest that for this kind of situation it is not impressive.

b) When extremes of all attributes are used, the impressiveness ranks high; Mode 10 is the highest of all in the final summary, as well as in all other tables and graphs except one. (Graph 2.)

c) The one notable exception to the superiority of Extreme Change over No Change is in the matter of Time, Modes 4 and 5. Mode 4, No Change of Time, tends to rank higher than Mode 5, Extreme change of Time. But conclusions based on this condition must not overlook certain characteristics of the audience and of the stimulus.

C. *Function of No Change in Attributes*

a) The three lowest modes are all of the type that eliminates change of one or all attributes of sound. The least impressive of all modes is that which eliminates all change, Mode 11.

D. *Functions of Each Attribute Separately Considered*

a) *Pitch.* Mode 2 (Extreme change of Pitch) shows a higher rank than Mode 3 (No change of Pitch) in Series I and in the total; but not in Series II. In Series I the difference is marked.

b) *Intensity.* Departure from Median Intensity shows low impressiveness (Modes 6 and 7), in the final summary.

c) *Quality.* Quality was isolated only in Series I; yet the results of Series I suggest that the dropping out of changes in Quality makes a pronounced difference in impressiveness; for the gap between Modes 8 and 9 is one of the longest in the whole body of graphs.

d) *Time.* Departures from median range in time makes little, if any, loss of impressiveness. Both No Change of Time (Mode 4) and Extreme Change of Time (Mode 5) show high rankings in the final summary.

II. Individual Differences in Auditors

A. *General Divergence*

a) Auditors vary in average total retentiveness between .15 and .87.

b) One subject, No. 1, stands out as markedly superior in retentiveness to all the others.

c) Four subjects, 11, 15, 21, 23, make a distinct class, which is low in retentiveness. The remaining auditors form a middle class of fairly distinct uniformity.

B. *Special Divergences*

a) The two auditors lowest in total reporting response are foreign-born, one a Filipino, No. 21, and the other a Chinese, No. 23. These two auditors respond to only a few modes. They agree so nearly on Mode 4 as to provoke the question whether this is not a mode commonly acceptable in Oriental speech, Uniformity of Time. Changes in Intensity seem to disconcert them (Table I, Modes 6 and 7), as also departures from Median change of Quality (Table I, Modes 8 and 9).

b) Subject 11 is worthy of special note. He reports that he almost never reads stories. Yet the results of a check test, calling for an immediate report, shows that his low record arises not because he does not apprehend the import of what is read; for he retains well under immediate report. In Series II he shows almost a 100% increase in response, due probably to training during the process of the experiment.

c) Subject 15 asserts that he lacks visual imagery. Whether so or not, he also shows learning in the process of the experiment. His earlier records are for the most part blanks; toward the end of Series II his records are relatively full.

d) Subject 1 is an exceptional case. His reports indicate that he is set to apprehend under all circumstances and that he retains well what he apprehends. The latest trials of the experiment indicate that he had become adapted to all modes of reading, no matter how foreign to his previous experience.

e) All women subjects—numbers 22, 24, 26, 27—rank well up in the total average.

f) The chief divergences in ability to report are shown by the male subjects, ranging as they do from .15 to .82.

Conclusions

1. There is a presumption in favor of using an extreme degree of change in all four of the attributes of sound during speech, especially for the purpose of securing retentiveness over an extended time.

2. The four attributes of sound differ in their effect upon the responses to oral reading.

3. Individuals differ widely in their responses to a given combination of change activities.

4. Auditors tend to become adjusted to unconventional modes of reading.

Subsidiary conclusions are:

a) In intellectual impression (such as this type is) regulation of Pitch changes is significant: a wide range is preferable to no change.

b) An even rate is superior to a rate excessively broken.

c) Interference with normal degree of change in Intensity reduces apprehension, retentiveness, or both.

d) Elimination of changes in Quality (Timbre) is greatly inferior to a wide degree of change.

General Conclusion as to Type of Experimentation. As to whether this method offers promise of valid results for further use, the following evidence is offered. (1) A judgment practically unanimous among students of oral expression is that changes in Pitch are always significant, that absence of all change is detrimental to intellectual impressiveness, and that the most effective reading is with an abundant, even extreme, range of change in all four elements of sound. This study brings results precisely in accord with these judgments. (2) Among authorities in elecution there seems to be no unified opinion as to the effect of changes in Time; and this study leaves that issue with results that are inconclusive, calling for refinement of method or more data. (3) Interference with changes in Intensity so as to affect accentuation, could be assumed empirically to be detrimental to impressiveness; and the experiment reveals just such a result. (4) Changes in quality are commonly assumed to represent changes in emotion, mood, total attitude, and hence highly impressive. This study strongly confirms this judgment.

Hence it is a justifiable conclusion that a method which brings results that conform with generally accepted judgments on matters of widespread experience, will prove useful in deciding ultimately such issues in this field on which opinion is not as yet in agreement.

BIBLIOGRAPHY

1. BEER, MAX. Die Abhängigkeit der Lesezeit von psychologischen und sprachlichen Faktoren. *Zeitschrift f. Psychologie,* 1900, 56, 264-298.
2. BLANTON, S. *Quart. Jour. of Public Speaking,* II., 4, pp. 352 " The Voice and Emotions."
3. GAUTHIOT. De l'accent et de la quantite en literanien. *La Parole,* 1900, II.
4. ROMANES, G. J. *Mental Evolution in Man.*
5. RUSH, JAMES, M. D. *The Philosophy of the Human Voice,* 4th ed., 1855.
6. RUTZ, OTTMER. *Sprache, Gesang, und Körperhaltung.* Munich, 1911.
7. SCRIPTURE, E. W. *The Elements of Experimental Phonetics,* 1902.
8. SHERIDAN, THOMAS. *A Course of Lectures on Elocution,* 1807.
9. STEELE, JOSHUA. *An Essay toward Establishing the Melody and Measure of Speech,* 1775.
10. ———, ———.*Prosodia Rationalis,* 1782.
11. WALKER, JOHN. *Elements of Elocution,* 1781.
12. WATT, H. J. *The Psychology of Sound,* 1917.

VOCATIONAL SELECTION FOR SPECIALIZED TASKS

A Study of Selective Tests for Hollerith-Machine Operatives

By Lawrence Marcus, Columbia University

A team of star baseball players can be ruined by the manager's failure to recognize their grooves. This is quite as true, even though less apparent to the eye, in industry where producing and selling teams are playing for efficiency. As an illustration of the possible evil effects of unwise selection let us choose the case of the Hollerith card-punchers employed in the auditing of money orders.

This work is of a simple nature. The operative punches the keys of a machine not unlike a simplified typewriter, which perforates cards according to the numbers on the cancelled money orders. Rapid co-ordination of eye and finger is the main requirement of a good worker. The operative will improve immensely with practice, and as the operation becomes automatic, the individual's output will increase greatly.

Let us compute the possible losses incurred by poorly selecting these operatives. Whether the workers are poorly selected by present methods or not will be discussed later. If 20 operatives are to be selected from a group of 100 candidates, according to hit-or-miss selection 10 "good" workers and 10 "poor" ones will be chosen By "poor" worker is here meant one who would come into the lower 50 if the group of 100 candidates were arranged in an ascending order of merit. Of the 80 rejected candidates there are about 40 who are better than 10 of those selected. Now let us see what the result of such a random selection is.

Assuming that a fair scale of pay for the work done by a "good" operative is $100 per month, the loss for the first two months on the average novice is $30. Because of the unfamiliarity of the work, the operative produces in the first month only 38 per cent of what she will produce when she has reached her practice level, and in the second month 52 per cent of her maximum. If the minimum entrance salary is $60 a month, and the output is 38 per cent of a hundred dollars worth, it is a simple problem in arithmetic to compute how much the organization loses on each new operative in the first two months of her service. $2\times60-(38+52)=\$30$. This is a necessary evil which is made up by the "good" operative later, for the pay adjustments necessarily lag a month or two behind her increases in efficiency. But in the case of the operative who will not later make "good," this is a complete loss, and since the individual practice curve is a varying, uncertain thing, there is no way of predicting before the eighth or tenth month whether any one girl is going to make a good operative or not.

This loss on the individual operative is not the only financial loss. The overhead of training, supervision, floor space and furniture cannot be computed so exactly, but conservative estimates point to a total of $11,600 saved on every 100 operatives. By proper selection a force of 100 workers can be reduced by 12, with the following approximated savings itemized:

Salaries (12 workers at $75 per month) annual..	$10,800
Supervision (Part service of a $1,600 supervisor for 50 operatives)...........................	400
Space (At 50 square feet per worker, and space selling at 50c. a foot)........................	300
General Overhead (Machines, light and heat, furniture, illness, etc.)......................	100
Total Savings..........................	$11,600

If we then note once more that the loss to the organization is $30 for each operative who does not make good, we can estimate in terms of money, just what random selection does for the careless employer.

There is another factor which is not so easily measurable in terms of money, but has a money equivalent nevertheless. There is a certain feeling of contentment possessed by the worker satisfied with his job and interested in it. When a man gets into the wrong job, he labors under a strain, both mental and physical. He is dissatisfied, frets, cannot bring himself to do his best work, and is much more likely to resign than the contented man who knows he is in the position for which he is best fitted. The physical strain naturally results from racing to keep up the pace and do a task set too high.

This strain causes over-fatigue, and the errors and decreased output that result from fatigue. That a healthy, contented work-force is an asset has been indisputably proved by the experience of those commercial houses that have installed rest-rooms, medical care, recreation rooms, etc.

To summarize, then, a poor selective process chooses for costly training many candidates who will never make good workers, and from this it follows that:

(1) Money is spent unnecessarily in the trial-and-error, hire-and-fire process;

(2) More workers are employed than would be necessary with proper selection;

(3) The general tone of the working force is lowered, offering fruitful soil for dissatisfaction and movements of unrest.

Various methods are used to-day by commercial organizations in selecting their employees. That they are making progress, even though very slowly, cannot be denied. The attention of business men was attracted by various sincere writers and investigators to the value of careful and scientific selection of their personnel, chief among these being Hollingworth, Gowin and Scott. But since business awoke to the need, there have risen a host of "guides," "scientific employment managers," "personnel engineers" and other pretenders who cry that they are ready to meet the demand with "scientific methods" and "tests." For the most part these claimants, prepared overnight to act as Applied Psychologists, are insincere charlatans, parasite camp-followers of the scientific bearers of an Idea. These quacks have recently come into the field to profit financially by selling business men bogus selective methods. The well-meaning and receptive business man usually buys the "expert's" services without inquiring very deeply into his *proven* methods, without asking for an analysis of the quality of the goods.

It appears that Applied Psychology is going in for a period of exploitation which must inevitably be followed by a period of skepticism and suspicion on the part of the business world. This has happened in the case of Efficiency systems. Taylor and Barth and Gantt were able and sincere men, devoted to the investigation of efficiency and best-method work, but these lions were followed by a drove of jackals and business was not able to discriminate. We now have the reaction. Conditions will doubtless be bettered, methods of judging the self-named "efficiency expert" will be found, and order will come out of the chaos. So with "Vocational guides," and "personnel managers," but only after the business man learns to discriminate.

Besides the physiognomist, there are the phrenologist and the palmist. These are not so successful in selling their wares. The palmist is generally discredited by the business man, but although he no longer believes in life lines, he still listens respectfully to orations on the "Significance of hand and fingers" by the physiognomist. The phrenologist capitalized his bumps for a time, but is fast losing the confidence of business executives. Proof is not forthcoming that a bump over the right ear is evidence of brains dedicated to salesmanship, and the phrenologist is bound the way of the clairvoyant and palmist.

There is yet another vocational "specialist." He invents "tests." Sometimes he calls these "psychological tests," but they are no more psychological than tests. Some of these tests are ingenious, others are quite simple, and their value lies in the confident statement of the inventor that they select the type of worker desired. If, for instance, a number of girls are to be chosen to operate certain machines, the candidates are placed before an innocent little board with holes in it. They are told to plug certain of these holes with matches. Upon their efficiency at this task is based the judgment as to whether or not they will be able to operate their particular machine skilfully. The assumption is as follows: motor co-ordination is required to run a machine; motor co-ordination is required to perform this match-hole test; therefore efficiency in one is guarantee of efficiency (at least potential efficiency) in the other. It is a bright thought, and so simple, like all great revolutionizing thoughts. But the over-curious student asks impersonally whether it has been proven, mathematically or in any way other than the inventor's statement, that a valid correlation exists between machine-operating and hole-stuffing? It is hoped that a correlation does exist, because it is so distressing to see confident but unsupported statements go into the scientific discard.

Of course, the great majority of business men have not called vocational doctors; nor have they swallowed sugar-coated but futile pills. They know that all is not well with the world, but they declare that they feel vocationally "so-so." The average business man advertises for a certain type of employee and reads in testimonials what his former employers have to say about his abilities.

The testimonial is a peculiar document. Were one to call the average executive's attention to the fact that he is "strongly exaggerating" Jones' virtues and thickly slurring his vices, he would be very indignant. The fact remains that Jones is not deserving of all the glorious garments that are draped on him. But the writer of the testimonial knows that a virtuous Jones must emerge if the testimonial is to be of any value, so he gets out his bright colors, sets himself squarely to admire Jones and daubs a few edifying adjectives. The words, unusual in ordinary business English, sound well and he grows bolder, adding finer and better sounding words and phrases, until he feels himself edified, sharing with Jones his glory. Thick and eloquent the phrases fall in his exultation, and Jones appears—brilliant.

The new employer knows the grand and glorious feeling of writing testimonials, so he discounts the exaggerations until an approximate Jones results. By his own letter and in an interview Jones tells what he has done in the past, and this is taken as prophecy of his future value. The larger corporations have application blanks in which the applicant tells about himself in a regulated way. Most of these blanks are designed to arrive at the candidate's experience briefly. Brevity seems desirable, so Jones is asked to state briefly on two short lines, just what he did for the past seven years. Some of the blanks ask questions which are intended to draw forth replies that form a picture of the candidate's character.

One blank asks a candidate whether he has "ambition," "mental accuracy" and "patience," among a dozen other traits, and in a space about 2 inches by 1.2 inch he is asked to tell why he thinks he possesses each attribute. The question was respectfully put to an executive who put much store in this blank as to whether in his opinion a shrewd applicant could not answer in the affirmative to all the virtues and give plausible reasons for each. He wagged his head in satisfied fashion and said, "We find out what we want to know from this," implying that someone on the staff could analyze the replies, and by properly weighing exaggerations and "reading between the lines" could lay bare the applicant's soul and get his true measure. If this is to be considered true, it should be proven.

Since neither the palmist, phrenologist, character-reader, "test"-specialist, nor vocational "guide" offers selective methods that stand up under analysis and requests for proof, what *is* the best way known thus far to select 10 girls for a specialized task from 30 that apply?

If, for example, a correlation is found between the ability to assemble a simple electric bell and the ability to repair motor cars, then the man who does well as a bell assembler will make a good auto mechanic. If the ability to typewrite is found to correlate with the ability to cancel the letter A on a page of reading matter, then a girl who cancels A's rapidly and accurately will make a good typist. Why not, then, choose your typists from these candidates who cancel A's well? A sample of the actual work can be obtained so easily in the case of typists that the value of the indicator here lies rather in its use as a vocational guide than as a vocational selector.

A typical procedure which illustrates the method and its possibilities is brought out in the following study of the tests used in the selection of Hollerith machine operatives. The present tests are analysed, certain empirical tests are applied and the results compared. The problem is to select from inexperienced female candidates those who will become rapid and accurate operatives of the Hollerith statistical machine. The Federal government requires the services of many of these operatives and they are chosen by means of the examinations of the Civil Service Commission.

According to the Civil Service Law of 1883:—

" Such examinations shall be practical, in their character, and so far as may be shall relate to those matters which will fairly test the relative capacity and fitness of the persons examined to discharge the duties of the service into which they seek to be appointed."

The Commission must, then, test the " relative capacity " of applicants for government service by means of an examination.

Let us examine the Civil Service tests with this question in mind: "Are these true indications of the applicants' potential efficiency on the job?"

It appears that excepting for technical specialists, such as chemists, engineers, etc., these tests are all scholastic ones; that is, they test the scholastic ability of the applicant—they determine how much learning the applicant has acquired. Whether scholastic ability is an indication of clerical or subclerical ability is an important question and in the interest of the team should be based on something more substantial than an assumption. The following inferences are made without sufficient foundation:

(1) That scholastic ability is a measure of native intelligence;

(2) That intelligence is a prime factor in the personal efficiency of subclerical or specialized workers (whose task may require rapid recognition or continuous rapid motor co-ordination);

(3) That this general native intelligence is equally an asset at all such tasks and that, for instance, an efficient money counter must also be a good press-feeder.

These inferences, which are founded on assumption rather than facts, are frequently untrue, and in these frequent cases, square pegs are badly fitted into round holes with the result that the Service suffers.

In order to correlate *efficiency as card-puncher with ability to score high on the Civil Service examination,* we must first establish our orders of merit. Let us arrange those who have already been tried out on the job in an order of merit, the best worker being number 1 and the poorest worker being 115.

Working in the Post Office are some 150 operatives whose output is recorded daily, both for quantity and accuracy. The total efficiency of any worker is the product of her speed times her accuracy, computed in the following way:

$$\text{Efficiency} = \text{Speed} \times \text{Accuracy}$$

(1) A daily output of 3,200 is considered 100% Speed.

(2) Total absence of errors is considered 100% Accuracy.

(3) For every error, 50 is deducted from the actual output of cards, with a proportionate drop in the speed rating.

(4) The accuracy per cent rating is obtained by dividing the "revised speed rating" by the "actual speed rating."

An example:

Miss X punches 2800 cards with 4 errors.

(S) $\text{Speed} = \frac{2800}{3200} = 87\frac{1}{2}\%$

(A) $\text{Accuracy} = \frac{2800 - 4(50)}{2800} = \frac{2600}{2800} = 93\%$

(E) $\text{Efficiency} = S \times A = .875 \text{ x } .93 = .814$

This worker's efficiency that day is 81.4%.

This is not the whole story, however, for in order to get a valid "order of merit" in card-punching efficiency we must take into consideration the effect of practice. That is, we cannot compare Miss Jones who has been in service two months with Miss Smith who has been in service for thirty-two months. In order to allow for practice, we plotted the practice curve, which we based on the records of twenty efficient operatives who had been in service for over two years and who apparently had passed their practice period. The procedure was as follows:

With 3,200 as the daily average of cards punched, and with a penalty of 50 cards for each error, the final efficiency of 20 "good" operatives was averaged and plotted. In the twenty-seventh month of service the 100 per cent mark is reached, i. e., either the operative punches 3,200 cards as a daily average without errors, or she punches so many more that her errors bring her down to a credit of 3,200. The units (per cents by months) are given here both as the monthly arithmetical average as corrected by "smoothing the curve."

PRACTICE CURVE OF 20 HOLLERITH CARD-PUNCHING OPERATIVES

Month	Per Cent Efficient		Month	Per Cent Efficient	
	Arithmetical Average	Smoothed Curve		Arithmetical Average	Smoothed Curve
1	38.28	38.3	15	94.25	94.4
2	52.43	51.9	16	95.72	95.1
3	60.96	61.4	17	95.33	95.7
4	68.90	68.3	18	96.39	96.1
5	73.89	74.2	19	96.83	96.7
6	79.74	79.3	20	95.45	97.2
7	82.95	83.4	21	97.14	97.7
8	87.31	86.3	22	98.81	98.1
9	87.44	88.4	23	98.38	98.6
10	89.71	90.0	24	98.62	99.0
11	91.62	91.1	25	97.86	99.4
12	91.51	92.0	26	97.33	99.8
13	93.14	92.7	27	98.52	100.0
14	92.92	93.5			

Consider an operative, Miss Y, who has been in service six months, whose record we have for the 73 days immediately preceding the eighth month allowing about 25 days for a working month. We compute the *normal* efficiency, that is, the average efficiency of the group, during the seventh, sixth, and fifth months (73 days' record) and compare with Miss Y's record. The normal efficiency is taken from the composite practice curve. The following proportion is true: The normal *perfected* output is to the normal three months' output as Miss Y's *perfected* output will be to her three months' average output. Then

$$\frac{100}{78} = \frac{x}{80}$$

Where 78 equals the average of normal outputs for 3 months (5th, 6th, and 7th), and where 80 equals Miss Y's average three months' output and where x equals Miss Y's efficiency after she has passed her practice period. By thus bringing the inexperienced operatives up to their theoretical final efficiency, we can establish a fair "order of efficiency."

The validity of this order of efficiency is discussed later when it is compared with 80 unrevised records, of 80 workers, each of whom had been employed for 20 months or more.

In drawing a practice-curve the two peaks, unusual jumps in the 13th and 24th months, appear at first glance to be caused by some factor entering those months and not the others. Nothing can be discovered, however, for there can be no seasonal element involved, since any worker may reach her 13th month of service in any seasonal month, depending on the date of her entry into service, and these are varied. Were it the 12th and 24th or 13th and 25th months, in which peaks occurred, the theory of an annual salary increase might be entertained. The peaks are probably no more than might be expected with the individual girls varying so in their course of practice. The peaks are well within the possible deviations. An inspection of the individual practice curves shows at a glance the lack of uniformity in individual improvement. A sharp rise in one, a weak catenary in another, etc., causes us to accept the peaks as not so unusual.

Having established our "order of efficiency," let us turn to our "Order of Civil Service Ratings." The examination as given by the Civil Service, takes about 2-3 hours, the candidates being assembled in one room and tested in the following subjects: Spelling

Arithmetic
Letter Writing
Penmanship
Copying.

The spelling test consists of 20 words, distinctly read and defined by the examiner. Arithmetic consists of simple problems involving fractions, decimals, per cents, etc. Letter writing is a test of writing (composition) ability, and copying consists of transcribing from plain and "corrected" copy. Penmanship is judged, if we can say that it is judged, on the merits of all the written work. A photograph is attached to the application blank which is required of every candidate.

When the examination papers are scored, certain weights are given the subjects. These weights are arbitrarily chosen, and apparently not very well chosen, for by trying various other weights a slightly higher correlation was found between the newly weighted totals and efficiency as a worker. See table later.

Now, if a coefficient of correlation of .5 should exist between CIVIL SERVICE grades and card punching EFFICIENCY, we could say that ability in one is an indication of the individual's ability in the other. A good card puncher will pass a good civil service examination, and conversely a person attaining a high grade in the examination will be a good worker on the Hollerith machines. On the other hand, should the correlation coefficient be less than .4, it is probable that only a chance relationship exists, and a high grade in the civil service examination gives no indication of what the worker will do at the task.

We found the Coefficient of Correlation between Efficiency and Civil Service =.31. The weights which the Civil Service Commission uses are:

Subject	*Weight*
Spelling	3
Arithmetic	5
Letter Writing	3
Penmanship	3
Copying	6
	20

Four other ways of weighting were tried by rescoring each paper, and the TOTALS then correlated with EFFICIENCY. The results are tabulated on the following page:

Subject	Weights Civil Service's Own	A	B	C	D
Spelling	3	4	1	1	1
Arithmetic	5	5	3	1	1
Letter Writing	3	3	3	1	1
Penmanship	3	1	0	1	0
Copying	6	2	2	1	1
Correlation Co-efficient (r)	.314	.311	.273	.343	.345
P. E. of r.	.06	.06	.06	.06	.06

It will be observed that the best result is obtained by leaving the subjects unweighted and discarding penmanship entirely, and if no better tests are found, it is recommended that weighting system D be adopted for these subjects. This recommendation is based on the results obtained by the examiners with the present method of grading penmanship. At present over 9/10 of the grades are included in the range 76%-80%, with 77% as the favorite mark. About ½ of the papers are scored 77%. This is not very discriminating.

It is further recommended that, if the Commission seeks to measure the handwriting ability of the operative, some standard method of grading handwriting be adopted. A standard scale serves to overcome the examiner's individual idiosyncrasies and gives the judgment an objective basis. And the Commission's attention is invited to the Thorndike Handwriting Scale. It has been in use for 10 years or more and is, therefore, comparatively new, and has probably not come to the Commission's attention, but in view of the results it has produced the Commission can safely venture to give this method a trial. A further suggestion as to the present examination is that the number of test-words given in "spelling" be doubled.

With these slight changes, i. e., handwriting and spelling, it is possible that the correlation coefficient may be brought up into the doubtful .4-.5 or even over .5 into the "certain-indication" level.

Correlations were computed between EFFICIENCY and EACH SUBJECT, and the following coefficients found:

Subject	Correlation with Efficiency	P. E.	
Spelling.........	r = .29	.06	
Arithmetic......	r = .30	.06	
Letter Writing...	r = .25	.06	in all these,
Penmanship.....	r = .14	.07	n = 115
Copying........	r = .21	.06	

Since the operation of the machine requires a certain nimbleness of finger, and since the ages ranged from 19 to 54, it was thought that a negative correlation between AGE and EFFICIENCY might be found, but no correlation resulted, r being —.003.

In search for a better indicator of card-punching ability, we examined the workers with 9 psychological tests, mainly of the Woodworth-Wells series.

1. Cancellation (Cancelling "A").
2. Opposite (Given a word, to write its opposite).
3. Whole-part (Given an object, to write some part of that object).
4. Substitution (100 geometric forms, random distribution, to write in each form a number according to a code).
5. Directions (The task being to obey complicated directions).
6. Number checking (To draw a line through every number containing both a 5 and a 4).
7. Completion (To fill in the words of mutilated sentences).
8. Verb-object (Given a verb, to write its object).
9. Logical Sequence (Given a paragraph with sentences disarranged, to put them in proper order).

Each test was printed on a different sheet.

The directions which were printed on the reverse of each sheet included examples of what would be found on the test paper. These directions were supplemented with verbal instructions until each operative understood what she was to do. At a signal the sheet was turned over and work begun.

Scoring

In all of these the time was fixed as indicated, the score being the number of correct reactions.

(1) Cancellation of A's. Time 60 seconds

Credit 2 for each one right, 1 for each one half-right (questionable reaction) and penalize 3 for each wrongly marked.

(2) Opposites. Time 90 seconds
Credit 2 for each one right, 1 for each one half-right (questionable reaction) and penalize 3 for each wrongly answered.

(3) Whole—Part. Time 50 seconds
Credit 2 for each one right, 1 for each one half-right (questionable reaction) and penalize 3 for each wrongly answered.

(4) Substitution Time 110 seconds
Add the number correctly done and penalize 2 for each wrongly marked.

(5) Hard directions. Time 90 seconds
Credit correct reactions and penalize 2 for each incorrect reaction.

(6) Number checking. Time 80 seconds
Credit 2 for each one right and penalize 2 for each one skipped and 3 for each wrongly marked.

(7) Completion. Time 360 seconds
A credit of 10 was given each sentence correctly filled out and partial credit was given on the basis of the number of omissions and the number of blanks to be filled out.

(8) Verb—Object. Time 50 seconds
Credit 2 for each one right, 1 for each one half-right (questionable reaction) and penalize 3 for each one wrongly answered.

(9) Logical sequence. Time 70 seconds
6 sentences in exact order equals 100; 5 sentences in exact order equals 80; 4 sentences in exact order equals 50; 3 sentences in exact order equals 20; less than 3 sentences in exact order equals 0.

The following introduction was given the experiment: "We are going to perform an experiment in Applied Psychology. Doubtless you have heard of Prof. Münsterberg and his psychological experiments in the movies. The purpose of his experiments was to test the power of observation. Just what character of the mind *we* are going to study, I am afraid I cannot reveal until the experiment has been completed. To do so would vitiate the results. This experiment is being done with the direction of the Auditor, and it will be to your advantage to do your part as well as possible in these tests that are to follow. Some of these tests may appear nonsensical, but may I remind you again that it is imperative that you follow directions carefully."

The tests were administered to 2 groups of about 75 in each group, in a large room with chair and desk provided for each person. Both groups were examined between 10.30 A. M. and noon, of a bright, sunny day, the temperature being 73° C, and the test conditions being on the whole, fair.

After scoring, correlations were computed between EFFICIENCY and EACH TEST separately. Then TEAMS of 3, 5, and 6 tests were correlated with EFFICIENCY, the teams comprising those tests which had separately given high coefficients. The standing of any worker in the "Team of 3" ranking is obtained by averaging her rank in each of the 3 tests included in the "Team of 3."

The coefficients follow:

	r	P. E.
(1) Cancellation..........	.019	.06
(2) Opposites.............	.154	.06
(3) Whole-part............	.119	.06
(4) Substitution..........	.236	.06
(5) Direction.............	.046	.06
(6) Number Checking.....	.094	.06
(7) Completion...........	.211	.05
(8) Verb-object...........	.031	.06
(9) Logical Sequence......	.316	.05
Team of 3 Tests (4) (7) (9)	.453	.05
Team of 3 (Weighted: Test 4 by 31.3......... Test 7 by 27.5.......... Test 9 by 41.2).........	.391	.05
Team of 5 Tests (2) (3) (4) (7) (9).................	.448	.05
Team of 6 (Weighted: Test (1) by 5/4......... Test (2) by 4/3......... Test (3) by 1........... Test (4) by 2........... Test (7) by 7/4......... Test (9) by 8/3.........	.473	.05

In order to check the method of allowing for practice as described some pages back, correlation coefficients were computed for EFFICIENCY and SEQUENCE of 80 cases having 27 months or more experience. These 80 had certainly reached their practice level, and when the coefficient thus obtained is compared with that obtained for 138 cases (some prorated for practice effect) these two coefficients are found to be nearly alike ($r = .319$ in one case and .316 in the other). This seems to uphold the scheme of proportions by which the individual operative's efficiency was adjusted for practice. Another similar check gave the same result, the "r's" being .277 and .236 when 80 experienced (27 months) and 134 mixed workers are correlated for SUBSTITUTION.

Theoretically, the coefficient should be higher when 80 unadjusted cases are used, and this is borne out in the very slightly increased coefficients. The individual practice-curves which go to make up the composite curve may rise far above or below the main curve for any month and false adjustments are not compensating. They must inevitably transpose individuals in the order, with accompanying decrease in coefficient. In this case, however, few such transpositions have been made as witnessed by the check of the 80 unadjusted cases.

Separate correlations were made for each test for group II, some 50 cases. No coefficients of significance were produced, the P. E.'s being as large as the coefficients in half the tests.

The danger of prophesying as to which test will be an indicator of any performance is brought out by the low coefficients of Cancellation and Number-checking. From a superficial analysis of the task it would appear that a quickness of recognition and a nimbleness of finger were prime requisites. A sort of rapid complex reaction would be essential. And it was believed that Cancellation and Number checking might be good tests. They are not, for neither an r of .019 nor of .094 can be said to be indicative. Had we proceeded like the match-stuffing vocational specialist previously mentioned, we would have declared them valid tests without trial.

A table of correlations and their probable errors follows:

CORRELATION OF EFFICIENCY WITH TESTS

		Both Groups		Groups II		
	Cases	r	P. E.	Group II 50–56 cases	P. E.	Cases
(1) Cancellation........	(125)	.019	.059	.120	.091	56
(2) Opposites...........	137	.154	.056	.021	.092	50
(3) Whole—part........	137	.119	.056	.210	.087	50
(4) Substitution........	134	.236	.055	.081	.092	50
(5) Directions..........	133	.048	.058	.103	.091	49
(6) Number Checking...	137	.094	.057	.386	.079	58
(7) Completion.........	138	.211	.054	.085	.092	50
(8) Verb—Object.......	136	.031	.058	.118	.091	49
(9) Logical Sequence	138	.316	.051	.219	.087	50
Tests:						
Team of 5 (134 cases) (2) (3) (4) (7) (9).		.448	.045			
Team of 3 (134 cases) (4) (7) (9)............		.453	.045			

Team of 3 (138 cases, weighted as described on page 199).
$r=.391$ P. E.=.05 (Both groups)
Team of 6 (138 cases, weighted as described on page 199).
$r=.373$ P. E.=.05 (Both groups)
80 cases having 27 months' experience or more:
Efficiency and Sequence $r=.319$ P. E.=.05
Efficiency and Substitution $r=.277$ P. E.=.055.

Since the "Team of 5" Psychological tests correlates .45 with EFFICIENCY at Hollerith operating, and since the Civil Service Commission's examination correlates only .31, it is believed that the "Team of 5" is a better selecting medium for the reasons that:

(1) A coefficient of .45 is well into the "indicative" zone, while .31 is below even the hazy "zone of suspicion."

(2) .45 is about 10 times its P. E., while .31 is about 5 times its P. E. In spite of the fact that 5 P. E. includes about 99% of a normal curve, it is felt that this is not a superfluous point.

(3) Taking the coefficients on their numerical values, the Tests' coefficient is about 50% greater than the Commission's.

(4) The "Team of 5" can be completed in 12 minutes and can be scored in 2, while the Civil Service test takes several hours to write and at least 10 minutes for careful examination.

Conclusion

It is by such procedure as has been followed here that selective media will be established. The work of sincere investigators is beginning to bear results. Already there are tests for the selection of Typists, Telephone-operators, Stenographers, Label-pasters, and a half-dozen others. The day of the quack is near. His eloquent but empty booming will be silenced by mathematically proven facts about vocational selection and guidance.

A CASE OF PHENOMENAL MEMORIZING BY A FEEBLE-MINDED NEGRO

By Hiram Byrd, University of Mississippi

Eugene Hoskins is his name. He lives at Oxford, Miss., a University place of about three thousand people. He is well known about town for his eccentricities, but more especially for his uncanny knowledge of dates. A bystander said to him: "I was married on the 8th of June, 1901." Without a moment's hesitation Eugene said: "Dat was Satu'day." Given the month, day and year, he will give the day of the week. He never fails, never hesitates. Vary it if you will by giving the year and month and asking what day of the month was the second Tuesday, or the fourth Friday—he answers just the same. It is one of the diversities among the university students to get old calendars and try him out. He is a never ending source of entertainment for them.

I have said he never fails. That is, so long as you stay within his limits, for he has limits. Go beyond that and he is at sea. He can't go back beyond 1901, and can't go forward beyond 1924. But during these 24 years success is 100%. It should be noted, however, that his limits have not always been so advanced. Mr. Harvey remembers when he could not go beyond 1920, and Eugene himself admitted to me that he is advancing his limits and hopes to reach 1925 by next year. Asked how he does it, Eugene says he can't tell you—that he doesn't know himself. So the impression has gone forth that it is a sort of supernatural gift.

Eugene says he was born in Tate County, Miss., September 10, 1896, which is probably correct—one can almost imagine that he remembers it. He is rather tall, regular build, and quite black—the type that is known in the race as "eight-rock"[1]

He has a habit of shrugging his—I started to say his shoulders, but I believe it is his whole back that he shrugs. It certainly gives that impression. He also has a habit of mumbling to himself, and often laughs right out.

---

[1] A pure blood African is, among some of the Southern negroes, called "eight-rock," while half-breeds are known as "ginger-breads" and lighter shades as "high-yellows."

The first time he ever went through his paces for me he held out his hand for money. I gave him a dime which he accepted, but he will rarely accept more than a nickel. He has been known to refuse a dollar and become indignant when the donor put it in his pocket. He would not touch it and insisted that it be taken out. He is fond of music and keeps a guitar, but he can only thrum. He used to beg one of the students in the University (Mr. Harvey) who plays a violin to play with him. I lately arranged with Mr. Harvey to do so, and proposed it to Gene, but he assured me he could not play, and would not engage to try. He stays with the Gambles, who furnish him food and clothes and a place to sleep, in return for which he delivers papers, gets in coal, feeds the pigs on the farm, and does little odds and ends about the place. He trusts the Gambles implicitly, but buys his own sugar for his coffee. He is variously called Eugene, Gene, James and Jim. He resents being called Jim and will not answer to it. When the grass was burning and was about to reach the Rogers house, some one called out to "Jim" to bring a pail of water. But he balked and refused to move. When asked about it later he said: "Da wan't talkin' to me—I ain't Jim." But when Mr. Gamble calls him Jim it amuses him—he takes no exception to it. He meets all trains that pass in the daytime and in the early part of the night. It is said that he meets them all. I had occasion to leave recently on a train between three and four o'clock in the morning. He was there to meet that train, although it was a very chilly morning. I was talking with him once at the Gambles, questioning him about himself and family, when the locomotive whistled. He darted out of the room almost like a released spring. I asked him yesterday if he meets all trains. He said not, and told me of two or three times that the trains had passed when he didn't meet them. (More light will be thrown on this curious passion for the trains later.) He has never been to school "to speak of" but keeps a pencil and a note book in which he prints out words. He spells out all the words he sees. In my office while taking the Binet test he stopped to spell out the word globe on a card index case.

When he first came under my notice I resolved to analyze the case and determine definitely the source of this unusual power. I sought all the information I could get from those who had known him best and longest, and this I supplemented with the Terman revision of the Binet test. It took several weeks making up to him to get close enough to him to make the test. The Gambles kindly helped me by telling him what a nice man I was and at the same time cautioned me not to call him "Jim." At length I struck his tender spot. After complimenting him on his unusual ability, I asked him why he didn't put it in the paper. Told him I would write it up for him if he wanted me to. "Den everybody see it," he said. But I didn't press the matter then. Only when I chanced to meet him I would manage to keep it before him. In the meantime I had the stage all set waiting for the psychological moment to test him. It came yesterday when I told him if we didn't do it during the holidays I couldn't take the time. He was then on his way to meet a train. Under suggestive pressure I piloted him up to the office and began the examination. He liked it and opened up freely.

There are several points of the examination that are worthy of note. The first is that his basal age is only four years. He failed at five to execute three commissions. From the basal age of four he "scattered" through to twelve, finally earning four additional years, making a mental age of eight. He did the most astonishing and contradictory things. For instance, although he has such a wonderful memory for dates, he failed utterly to repeat five digits in the seven year test. In the ball-and-field test he made a superior record, and in the Healy-Fernald puzzle succeeded in less than a half minute. But he could interpret a picture only in terms of enumeration, could define in terms of use only, and thought that *house* rhymed with *day*. He read with some hesitation, but with fair understanding, but only got four out of the twenty-two memories!!! His word test was interesting. He started off with names of places which I had to interdict, for as will be seen later, he has committed to memory a large array of these and could rattle them off a very long time before he would ultimately run down. So I suggested *table* as a starter, which he followed up with *chair* and *university* and then back to names of places. He was stopped again and started off with *hat*. From this he went to shoe, heel, shoe-shop, court-house, North Street, South Street, and here he stalled again. (The shoe-shop in Oxford is near the court-house, and North Street and South Street both lead directly from it.) He was encouraged and finally started off with the word *peanut*. He followed this up with corn, wheat, rye, hogs, cows, calf, bull, cat, dog, rat, mouse, lice, picnic, churches, and here he made his final stand. It is easy to see why one brought up on a farm would pass from corn, wheat, and rye to hogs, and on down to lice, but it would be interesting to know the historical setting that caused lice to suggest picnic to his mind.

It is worthy of note that in the test everything pertaining to ages and dates was perfect. Giving him full credit for this, his mental age would check up nine years.

The next question is how to account for his remarkable ability to name dates. The following facts will throw some light upon it. He allowed me to keep his notebook above alluded to. Printed out on the first page is: "Mississippi Division, Jackson District." Then follows the names of stations leading from Jackson, Tenn.—(Just as they were spelled out in his book) Bemis, Melases, Medon, Teag, Toone, Shandy, Bolivar, Hickery Valley, Temper, Grand Junction, Michigan City, Hudsonville, Holly Springs, etc. He has two pages of these names, aggregating 34, which he took great delight in naming off to me.

Then comes a page with the following words on it: "Engine Number is 1746 Run from Frogmoor to Mounds. Engine Number is 795. 919 914 906 851 945 887 Run from Water Valley to Frogmoor 1 Miles to Jackson." Given a start on this page, he reels off everything on it. Then comes another page which starts: "Northern Line Passenger Engine Number is 1140 1139 1008 1051 1108 1065 1080 1141. Run from Champaign to Centralia 130 Miles Illinois Division Champaign district."

He can repeat this page if given three or four words at the start. Several other pages of like matter follow, all of which he can repeat when given a start. Only the last page he can't repeat—it is too new. He has not committed it to memory yet. All this explains his passion for meeting the trains. It also explains that his method is one of committing to memory. Now add to this the fact that he is gradually extending the time over which he can give dates, and his method is at once understood. He admits that he doesn't always write his dates down to commit them to memory. In like manner he can repeat a large number of counties, with their county sites, and a large number of states with their capitals, but without order, and no group is complete. The idea of completion seems not to have occurred to him.

Conclusions

1. That his mental age is 8 to 9 years.
2. That his auditory memory span is rather short, failing on 5 digits.
3. That he has by effort committed to memory a large array of facts along three lines:
   a. Dates
   b. Places
   c. Locomotive engine numbers
4. That he is continuing his studies along these lines, and it may be predicted that he will ultimately become even more proficient.

OPPORTUNITIES FOR COLLEGE GRADUATES IN PSYCHOLOGICAL EXAMINING IN SOCIAL SERVICE WORK AND EDUCATION[1]

By MILDRED FRANCES BAXTER, Psychol. Examiner, Dept. Education, Rochester, N. Y.

The profession of psychological examining is still a comparatively new one—very new when we consider how long those of teaching, nursing, law and medicine have been open to women. However, the field is extending rapidly. This is perhaps one of the benefits derived from the war, for in no other way could this science have touched and become a vital factor in the lives of millions, in so short a time.

Because of the youth and rapid development of this profession, the requirements for it have not been generally standardized. Nevertheless, the strong belief that stress must be laid on the absolute necessity for adequate training is held by all who are working in the field. The best way to bring this home to you, I think, is to quote from a few of those universally acknowledged as leaders in the work.

Dr. Goddard of the Ohio State Bureau of Juvenile Research says, "The first requirement is a good foundation in practical psychology. A mere knowledge of terminology or book psychology is of little value, but one who has done experimental work, or any form of applied psychology is ready to begin. Just at present it is true that the demand is stronger for persons who have had sufficient experience to be able to go out and take charge, either of the examining in the schools or in an institution or some social organization. Yet the time is coming very rapidly when there will be openings for assistants in these positions and then much less training will be required. The great requirement, as always, is plenty of initiative, interest in the problem, and, of course, a good degree of natural ability."

Dr. Healy of the Judge Baker Foundation in Boston writes, "I do insist that the proper attitude toward this work is to look at it from a truly professional standpoint and acquire a professional rating that will place the whole work upon a dignified plain, where it should be. I think, then, the training demands post graduate work and particularly the degree of Ph.D. Perhaps you already know of the association of Clinical Psychologists; a doctorate is necessary for membership in this. This association has already been recognized in the State Law of California. The demand has been so great that much less well equipped women have been given positions, but in the long run the situation will be the same as it has been in medicine. In fact there is no reason why the study of the human mind should not be worth as much as the study of the body."

---

[1] Read at the Vocational Conference, Vassar College, Feb. 21, 1920.

Dr. Elizabeth Woods who was formerly a member of the department of psychology here at Vassar and is now Clinical Psychologist and Supervisor of Exceptional Classes for the state of Wisconsin, says, "I feel very keenly that there is need for a longer training for this work than many people realize when their interest in it is first aroused. Already considerable damage has been done by 'six-weeks' psychiatrists,' as some of the physicians in the army called some of the examiners. It can certainly be stated without reservation that clinical psychologists who have had the training which will qualify them for membership in the American Association of Clinical Psychologists are in demand and are getting salaries up to five and six thousand dollars a year."

The demand for pre-professional training for this work is being met by an ever increasing number of the Universities, Normal Schools, institutions and organizations. Some of the Universities now giving courses helpful in this line are—Harvard, Columbia, Carnegie Institute of Technology, Chicago, Leland Stanford, University of California, University of Michigan, University of Minnesota, Ohio State, Ohio University at Athens, University of Wisconsin and the University of Illinois. The course offered at Harvard during the summer is an excellent introduction to clinical work. Dr. Healy states "in Columbia, undoubtedly the best post graduate training is given."

Work with Dr. Healy of the Judge Baker Foundation, of which the psychological clinic for the Boston Juvenile Court is a part, is allowed to count for a Ph.D. at Columbia. This Foundation controls two scholarships, as yet he does not know what will be the situation in regard to them for next year. One of these is in an industrial school settlement—

for the better adjustment of those attending the classes, through thorough examining, oversight and following up. The Ohio State Bureau of Juvenile Research, which I shall mention again later, is taking on a few research students. This is an excellent opportunity for those particularly interested in juvenile delinquency. The Psychopathic Department of the Massachusetts State Hospital, in Boston, under Dr. Corwell offers internships each year to one or two girls with a good background of psychology. Maintenance is given for half time service, the other half may be devoted to study or a paid position. This experience does not necessarily lead to a position, it is rather one step toward a Ph.D. Exceptionally well prepared persons can occasionally get in with Dr. Adolph Meyer in the clinic in connection with Johns Hopkins University, especially if they are looking toward the study of medicine in connection with it. The Iowa Child Welfare Research Station under Dr. Bird Baldwin, established for the study of normal children, also takes on students. Dr. Woods suggests that many college girls will get an excellent start toward some sort of higher specialization by taking a summer school special teachers course and teaching in one or several types of the special classes such as have been organized for genuine subnormals, adjustment classes for pedagogical retardates and for gifted children. Dr. Woods will be glad to have any who might consider teaching in special classes communicate with her about positions in Wisconsin. The New York State Normal Schools at Oswego and Geneseo offer special training for the teaching of mental defectives.

The qualifications and requirements for a clinical psychologist will become more and more rigid. Already New York state has passed a law that a psychologist, to qualify for a certificate which will authorize him to examine for commitment to an institution, must have had at least two years graduate training and a certain amount of clinical experience. My advice to you is to get in all the post graduate academic work you can as soon as possible. You will need it and will want it. Once you are in the swing of the work you will not wish to step out and get the further training you wish you had.

The work of the psychological examiner as outlined in the report on the qualifications of psychological examiners, prepared by a committee of the American Psychological Association is defined to include the following: (a) the measurement of general intelligence and specific native abilities; (b) the measurement of educational attainment and vocational skill; (c) the collection of personal, social and psychological data regarding normal, defective, or gifted individuals; (d) the detection of the presence of the common varieties of abnormal mental conditions due to heredity, training or disease; (e) the giving of expert advice regarding the educational treatment of both normal and abnormal individuals.

This work is carried on chiefly by Departments of Public Instruction, by courts, in institutions for delinquents and for defectives, in industrial establishments, and to some extent by Boards of Health and Social organizations. The field ranges from a national bureau through organizations and institutions in sections of the country, in states and in cities, to private societies. There is the National Research Council which is comprised of eminent psychologists who were organizers of the psychological work done in the army. This council is issuing two sets of group intelligence tests which will be ready in the fall, these have been their main work for the past year. They hope to put out new sets each year for a period of ten years.

One piece of work that is covering the eastern section of the country is that carried on in the institutions for the blind under the direction of Professor Hayes of Mt. Holyoke College. Perkins Institute for the Blind, near Boston, is headquarters for the work in the north; the psychologist from this school goes to the schools in Hartford, Connecticut,—and in Batavia, New York. Overbrook in Philadelphia is the center for the more southern section; the psychologist in charge here goes to the Virginia school and the Louisville, Kentucky one, her assistant to Baltimore and Pittsburgh. At present the assistant's salary is $450 and all living expenses, the other two are paid $500 and $600. Within the next two years or so the western schools are hoping to employ psychologists. It is probable that these positions will pay better rather soon.

Several states have been alive for sometime to the problem of feeble-mindedness and are endeavoring to work out plans to meet it. The Massachusetts Society for Mental Hygiene has published a pamphlet containing a list of 39 out-patient clinics conducted in that state for the examination and treatment of patients with mental disease or defect. The cities and towns in the state which have psychological clinics are listed alphabetically with the name of the clinic, the location, the history, and information as to the special functions of each particular clinic. The state program, suggested by Dr. W. E.

Fernald, requires that every institution and school for the feeble-minded should conduct out-patient clinics, at the institution and in the various cities and towns served by the school; and that a travelling psychologist should visit at least once a year all rural districts not otherwise reached. The New York State Commission for Mental Defectives, in a report to the governor and 1919 legislative assembly, proposed the statewide establishment of clinics for the diagnosis and treatment of mental defect and disease. These clinics would co-operate with the health centers which the Department of Health is to establish throughout the state. Plans are now underway for the establishing of these by the commission. Emphasis is to be laid on mental defect as there are already clinics for mental disease.

One thing that points to an immediate extension of the field for psychological examining is the request that comes from the state assemblies for surveys of their territories. Such a survey was conducted in Montana a short time ago. One has been going on in Indiana for some time with the immediate result that during 1919 a law was enacted providing for a new institution in the southern part of the state. The Georgia General Assembly requested a survey which was conducted by Dr. Victor Anderson who, with an assistant, was sent as a scientific adviser by the National Committee for Mental Hygiene. In his report made in Oct. 1919 he recommends, "the creation of clinics manned by the staffs of the State Hospital for Insane and the School for Feeble-Minded to act as a Clearing House for the defective, subnormal, the peculiar and nervous children of the public schools, and the abnormal and delinquent children of the juvenile court and for various complex mental problems of the home and community." Other states have had these surveys and it is certain that still others will follow their example; it is likely that there may be a need for assistants in this work.

Many states are building new schools for the feeble-minded and it is very probable that clinics will be established in most of these. Delaware is building one in Sussex County near Selbyville; California one at Pomona and Massachusetts one at Belchertown. Tennessee has provided for the Tennessee Home and Training School for Feeble-Minded Persons; a provision in the law requires a mental examination within the first month of residence, a second within the first two years and a third within the first five years. The national school authorities having jurisdiction over federal districts and territories has added to their budget; if this sum was approved it would be used in part to found a School for Backward and Feeble-Minded Children in Mexico. A school is also being founded in the Hawaiian Territory.

In the District of Columbia a House Resolution recently introduced in Congress has provided for a Division of the Bureau of Education for the study of the needs of mentally handicapped school children. There are to be three branches —first, for the investigation and publication of the test results and all information along this line; second, a laboratory of mental tests and standards of mental capacity for normal children of different ages,—these essentially for the diagnosis and classification of mentally subnormal and abnormal school children, and juvenile and adult delinquents; third, a psycho-educational clinic for the examination and classification of school children from the District of Columbia or elsewhere.

Other states are leaders along educational lines. The Wisconsin State Department of Public Instruction has a State Clinical Psychologist and Supervisor of Exceptional Classes—whose duty it is—to diagnose problem cases in various cities throughout the state and to make recommendations for special classes. Besides being responsible for the personnel of these classes she must help in their organization and in the securing of a teacher for them, also so far as time permits she supervises the instruction. It is probable that this department will be enlarged in the near future. New York also has a state psychologist—Dr. Cornell holds this position with headquarters at Albany.

In at least one state—Michigan—the State Board of Health employs—in their Psychopathic Department—a psychologist and an assistant who is called a field investigator. This state gives mental examinations to all venereal carriers apprehended and treated by the state; it is the duty of these two persons to examine the patients in the hospitals throughout the state and to aid in obtaining social histories. The salary offered the assistant was $1,200 and expenses.

Again, the states are coming to appreciate the necessity of knowing the mental status of the children and adults brought before their courts and of the inmates of their reformatories and penitentiaries in order that they may handle the cases to the best advantage of the individual and the state.

Ohio has been a pioneer state in what seems to be a most important phase of the work. Nearly two years ago Dr. H. H. Goddard who for many years has been director of the research laboratory at the Vineland Training School for Feeble-Minded in New Jersey went to Ohio to begin the work of the State Bureau for Juvenile Research. Any judge in the state of Ohio may send any juvenile case to this bureau for an indefinite period of observation. After a most careful study of the case from every possible angle—medical, psychological, social, educational, etc., the Bureau decides what appears to be for the child's best—whether he should be sent to an institution, if so, which one,—or whether he should be returned to his own home or perhaps sent to some other better home. This report is given to the judge and he must act accordingly once the case has been turned over to this department. The work of the Bureau is not confined to court cases however though these must be attended to first. It is also a consultant and advisory body for parents who have children who are educational or social misfits for reasons other than inferior mental ability.

Illinois is also contemplating a Bureau for Juvenile Research.

The Iowa Child Welfare Station for the study of normal children has already been mentioned.

As for the state penal institutions, clinics have already been established in Sing Sing and Bedford Hills, New York, also in other states. California introduced a bill early last year for a department of psychiatry and sociology in the state penitentiary at San Quentin. Pennsylvania is considering the mental examination for classification of inmates of the prisons and reformatories. Mrs. Falconer, formerly at Sleighton Farm, the girls' reformatory for eastern Pennsylvania, realized the need long ago so that institution has had a resident psychologist for at least seven or eight years. The duties of the psychologist in this institution are to examine all who are committed to it, to report special examinations to the courts, to advise in matters of discipline, to aid in the classification of the girls in the dormitories and in their academic and vocational training within the institution and in the selection of those who should be given special school opportunities outside the institution which may in some cases lead to a college education. The salary there has been maintenance and $50 a month which is estimated as the equivalent of $1,200. Experience of this kind is invaluable to anyone who aims at court work.

Of New Jersey, Dr. Edgar A. Doll, Psychologist for the New Jersey State Department of Institutions and Agencies writes, "the Department of Institutions and Agencies controls all the institutions supported by the state in whole or in part. This Department has definitely planned a Division of Medicine and Psychiatry. When this Division becomes fully functioning it will undoubtedly have a section for psychology. The purposes of this section will be to afford psychological service to the larger State institutions. The functions of the institutional psychologist will be to make surveys of the mental capacities of the institutional inmates with a view to transferring the inmates to the appropriate institution or to provide for their proper development or placement in the various institutional activities. Undoubtedly, when this program is fully instituted, there will be positions in the State of New Jersey for psychological examiners in the form of social service work. At present this program is being partially carried out by means of a psychiatric clinic under the direction of the State Hospital at Trenton. The plans of the psychiatric clinic may later call for a woman psychologist in one of the correctional institutions. This position is not, however, as yet definitely created. When created it will probably pay about $1,800 and its duties will call for group and individual examinations leading to adequate psychological basis of classification and placement."

We have considered the situation where the state is the unit; in at least one instance a county is the unit. Erie Co., New York, has a psychological clinic conducted at Buffalo by the Children's Aid Society and the Society for the Prevention of Cruelty to Children. The children cared for at the Shelter, a temporary home, are given a thorough physical and mental examination. The psychologist's diagnosis and prognosis are considered a very important item in deciding the next step for the child's future, especially in cases which might be placed out for adoption. At present one experienced psychologist is employed by these organizations.

When we come to consider cities as the field for work we meet the same types of problems to quite an extent. In education we find that the Boards of Public Instruction of all large cities have a Child Study Department or a Department of Research. The purpose of this branch is usually thought to be the weeding out of the subnormal children from the regular classes. In part, of course, it is; they may receive a far more helpful and for them practical training in special, vocational or shop classes, and the brighter children will profit by the removal of the dull ones. On the other hand, there are pupils at the other end of the scale who are so alert and capable mentally that they too present a problem for the school administrator in as much as the organization of the public school is such that it fits about 60-70% of the children. These departments may also aid in the organization of an entire school, for instance, Junior High Schools where large groups of pupils come from the sixth grades of several different Grammar Schools in different localities embracing children from various grades of American and foreign homes. By means of group intelligence tests which supplement the teachers' judgments these pupils are classified according to their general intelligence.

The largest cities have for some years had clinics in connection with their courts,—especially with their juvenile courts. Chicago's is known as the Juvenile Psychopathic Institute and is under the direction of Dr. Herman Adler. Besides more experienced psychologists he has assistants who do routine examining but are not permitted to interpret the results. The Judge Baker Foundation under Dr. William Healy who began the Chicago work studies the problem of juvenile delinquency in Boston. Occasionally there is an opportunity for persons to receive an excellent training here. There has been a clinic for the Municipal Court in Boston under Dr. Victor Anderson, he has left and whether it is still in operation I do not know. The Psychopathic Clinic of the City Children's Court of New York has two branches, the Manhattan Clinic and the Brooklyn Clinic, each employs one psychologist. Philadelphia also has some provision for psychological examining for the courts.

Probably a new endeavor in the field of public health service has been undertaken by the Newark, New Jersey, Department of Health. In 1908 it had established psychopathic wards in the City Hospital to care for mentally afflicted persons pending trial to abolish the practice of incarcerating them in cells. In 1907 a clinic for the treatment of nervous and mental diseases had been opened at the City Dispensary. Experience from the work of these institutions showed that the Department of Health was the logical center for the coordination of all activities pertaining to mental health. Among the activities of the Bureau will be the care and treatment of mental disorders and maladjustments of behavior, the examination of juvenile offenders referred by the Children's Court, the mental rehabilitation of discharged soldiers, the examining of mental cases from the Child Hygiene and Venereal Disease

Bureaus, the collection of statistics as to the number of feebleminded, epileptics and delinquents in the community and the spreading of practical knowledge of preventive measures. In Newark also there is Dr. Maxfield's clinic in connection with the schools. It is possible that persons of relatively little experience may find some opportunity for work with him.

Private agencies and societies are coming to recognize the need for psychological advice, particularly those interested chiefly in children. In Boston the Children's Aid Society and the New England Home for Little Wanderers give routine tests; the Church Home Society in that city has adopted the practice of routine psychiatrical examinations for all children admitted to its care. They have a psychiatrist on their staff who examines, observes and treats the children.

In New York City there are several clinics, the directors of these may have need of less experienced assistants. Dr. Schapp has a clinic, which was formerly known as the Clearing House for Mental Defectives, held at the Post Graduate Hospital. He makes examinations of all kinds of cases from babies to adults and wishes to know their mental age, where possible, before he examines them. On Saturday mornings he conducts a lecture and gives demonstrations for doctors. He employs one psychologist for part time work for which he pays $50 a month. Miss Irvin on 64th St. conducts examinations for all children about to enter any grade in the 64th St. school, and they are classified accordingly. This is not financed by the Public Schools. Volunteers who are familiar with tests may find opportunity for experience here. Miss Farrell has one in connection with the Special Classes of the city schools. Dr. Heckman also, has one at New York University, but it is not likely that he has any call for volunteers.

Louisville, Kentucky, has a clinic which stresses the intimate relationship between organized social service and mental hygiene, under the joint control of the Board of Education and the Louisville Welfare League. Under the latter are federated the social agencies of the city. These agencies refer cases in which mental trouble is suspected to the clinic. The clinic is at the service of the juvenile and police courts; in the latter it is a matter of routine for this clinic to examine all infected women quarantined by the venereal clinic. Surveys are being made of institutions such as orphanages, and reformatories for the identification of mental defectives. It examines all admissions to the Kentucky Children's Home

Society. It assists in placing a check on the public school classes for superior children, "opportunity classes," classes for the retarded and the definitely defective. A mental survey of the Louisville schools is a project under the direction of the clinic.

The trend of the times in business and industry is to lay more stress upon the need for psychological selection of personnel and psychological direction of welfare work. These positions are ordinarily paying from $150 to $200 a month. The duties of these positions will particularly call for knowledge of sociological relations, knowledge of educational measurements, knowledge of group psychological tests, knowledge of clinical psychological tests, ability to estimate vocational aptitudes and industrial capabilities, and particularly ability psychologically to analyze personality make-up. Miss Rachel Pflaum, director of the Bureau of Occupations for Trained Women in Philadelphia, in an article in the *New York Evening Post* for February 5th states that, "more and more the large industries are coming to recognize the fact that women are better fitted than men for some of the executive positions. This was proved in the shops within six months after the United States got into the war; proved, that is, in the shops that had not learned it through experience." In many industries women appear to be better fitted for the positions mentioned above. In some concerns they are known as Employment Managers, in others as Directors of Social Work. A few progressive concerns have also established research laboratories for the purpose of working up tests to aid in the selection of workers. The United States Rubber Co. in Hartford, Connecticut, has such a laboratory, also I believe, the Eastman Kodak Co. in Rochester, New York. The Black Co., in the clothing industry, in Cleveland, gives intelligence tests to its employees; this and other companies in Cleveland promote, to a great extent, on a basis of these tests.

The field of psychological examining is a most interesting one, it owes part of its fascination to its youth for we may hope to keep pace with it and to help it grow. The value of this work is being realized more universally all the time but still there is need of enlightenment on the part of the public and particularly on the part of the officials to whom we must turn for legal action. All is not plain sailing yet. Less than five months ago in a large middle-western city, a feeble-minded woman who had been a public nuisance for some time was brought before the court commissioner. An earnest appeal was made by a member of a social agency to have this woman committed to an institution on the grounds that because of her firmly established unsocial habits combined with her feeble-mindedness—she was a public nuisance. The official's reply was, "I should like to know what woman in the city of —— is *not* a public nuisance." The woman was not committed. About the same time this official was induced to commit a young girl to the state institution for the feeble-minded. There was a delay in carrying out the order and a week later this same official, though it was contrary to the law, married this girl and permitted her to leave the state. However, we will hope that this is an extreme case and take comfort in Dr. Samuel McCord Crother's statement that, "You can't educate a grasshopper. He's too busy hopping. The peculiarity of man is that sometimes you can induce him to stop and think."

INTELLIGENCE AT SENESCENCE

By M. F. Beeson, Ph. D., Colorado State Teachers College

It is generally assumed that deterioration of the mental processes characteristic of senescence begins at about the age of 60 or 65. There is a utilitarian reason for knowing just what changes in the mental processes take place at senescence, and at about what age these changes make their appearance.

In order to get some information on this subject, the writer spent five and a half days, from September 8 to 13, 1919, in testing twenty cases between the ages of 55 and 93 by the Stanford Revision of the Binet-Simon Scale. Ten of the subjects were men and ten were women. The average age of the group was 75 years.

The tests were given to unselected inmates of a Home for the Aged. This institution is a private foundation, and is very strict with regard to the requirements for admission. It requires that the candidates for admission be worthy individuals who have been good and industrious citizens. There is a long waiting list, so that careful selection among the applicants is possible. The inmates occupy splendidly equipped little cottages, either a married couple and one single man in a cottage, or four single persons in a cottage. The married couples keep house, and are supplied with all necessities upon requisition, and the unmarried ones take their meals at a central modern cafeteria. So the inmates of this Home are highly selected, and are not comparable with those of poor houses, county farms, or county, state, or municipal institutions.

On the other hand these inmates are probably not strictly comparable with old people of the same ages outside of the Homes. Their presence at the Home indicates that they have not generally been successful in life. Some, not lacking in mental capacity, are so deficient in certain traits of character and personality as to render them unable to compete successfully with the usual business or professional or laboring man. A few are victims of drink or other forms of dissipation. Some others are, according to the testimony of the psychiatrist in charge of the institution, lacking in sanity. One inmate is avowedly demented. A few suffer from paranoia. One spends much of his time sitting in the lobby of the most fashionable hotel in the city, smoking cigars with which the Home has provided him, writing letters on the hotel stationery, and talking to the guests about famous friends and acquaintances. One of the most intelligent of them has an uncontrollable propensity to beg, notwithstanding trials by the police court for the offense and frequent warnings and threats by the director of the institution. Another lives in filth, though placed in the most sanitary and healthful surroundings. Several of the old men show signs of abnormal sexual tendencies.

The difficulties of testing old people were found to be far greater than those involved in testing children or college students. In the first place it was necessary to explain to each individual briefly the purpose of the test. Usually a brief explanation was given to the effect that the purpose of the test was to discover what the mental activity of old people is in comparison with that of young people. It was also necessary to add that it was not to be a test of education. The general plan was then to tell them that the test could be illustrated by an example, and if they did not object to that question, it would be continued. Then some interesting test would be given to begin with. From that point on curiosity would usually prevail over their caution, and the test would be continued with no great difficulty. Another means of arousing their interest was to ask them the questions for filling the blanks on the first page of the Stanford Record Booklet, questions concerning their family history and education.

After the test was started, other difficulties were the desire of the subjects to wander away from the task and to talk about irrelevant matters, to ruminate, to give advice, or to discuss the questions. In most subjects the greatest difficulty of all was their garrulousness. With other subjects the difficulty was to keep them interested in the task.

In one or two cases it was necessary to complete the test in two sittings. In one other case the subject, after taking one of two of the tests somewhat against his will, refused to go further with the task, giving headache as the reason for not desiring to continue the work. Two men could not be persuaded to take the test, but all the rest consented after a little persuasion.

The average time required for the actual testing was one hour and fifteen minutes. In twelve cases, owing to the lack of time or the great amount of irrelevant conversation by certain subjects, it was necessary to use the abbreviated scale or to give the complete scale only for the simpler questions. Eight were tested with the complete scale. The average time required for testing these was one hour and thirty-two minutes, as compared with one hour and five minutes for the twelve tested by the abbreviated scale. All of the striking reactions were recorded, such as nervousness, lack of self-confidence, and attitude towards the test. The responses were recorded verbatim. Information was also secured concerning *their professions, interests, manner of spending their time now*, and other facts of general importance. Many of the subjects had visual defects. Following is a brief sketch of each case. The scores for the vocabulary tests are given in all cases where the subjects passed the average adult or superior adult vocabulary test.

1. (Complete scale.) Man, age 68; mental age (M. A.) 8-6; I. Q. 53. About 8th grade. Swede. Carpenter. Came to America 1871. Speaks English well. In Home 2½ years. *Very lacking in self-confidence; hesitant and uncertain; very* deferential. Realized his poor responses, and in spite of constant encouragement, often said, " No, that's very poor, but I could do better if I knew English better." Also made excuses because of "poor schooling." In the ball and field test said, " I don't understand that; I don't know nothing about playing ball." One of the most neatly kept of all the cottages.

2. (Abbreviated scale.) Woman, wife of the former, age 55; M. A. 10-1; I. Q. 63. Dane. Went to school until 12 years of age, then took a course in nursing in a hospital in Denmark. Remained in the hospital five years. One of the *neatest housekeepers in the Home. Very lacking in self-confidence*. Almost wept from embarrassment during the test. She put the examiner off twice when he came to test her, and attempted to do so a third time, but was always deferential and polite. Very apologetic.

3. (Complete scale.) Man, age 82; M. A. 10-10; I. Q. 68. Education to about 8th grade. In Home 2 years. Carpenter and "contractor." Former captain in Civil War. Difficult to hold to task. Wanted to talk, to philosophize, to give advice. In the test of naming 60 words, could not name isolated words, but invariably made random sentences. Very *self-confident. Garrulous. Said he would enjoy taking the*

test every day. Is in love with one of the old ladies in the Home.

4. (Abbreviated scale.) Man, age 73; M. A. 11-6; I. Q. 72. Attended school until 15. Was then in "sixth reader." Favorite subject, history. Carpenter, cobbler, butcher, miner, etc. In Home 1 year and 3 months. Spends his time reading, tending lawn, and trimming trees. Has seven daughters and two sons, youngest 17, oldest about 36. Has been ill a good deal—catarrh, hay fever, hernia, rheumatism, sun stroke at age of 19. Nervous. Trouble with eyes. Mind wandered from task considerably. Interested in birds; studies a bird-guide. Talks constantly, especially about other inmates. An example of the loose answers given by many is his reply to the question, "Why should we judge a person more by his actions than by his words?" The response was, "His outward appearance sometimes gives him away. Sometimes these confident fellows are dressed up well, and their actions would be fine. By George, they're just after the almighty dollar." In answer to the question, "What ought you to do before undertaking something very important?" the reply was, "Well, if that ain't another one. Very important—I'd be kinda dumbfounded on that."

5. (Abbreviated scales.) Man, age 79; M. A. 11-6; I. Q. 72. Education—sophomore in college. Occupation, mining. Born in New York State; 54 years in the West. Three and a half years in the Home. Suffers from asthma and leakage of the heart. Spends time reading daily papers and Literary Digest. Said by director of institution to be a case of paranoia. Goes to fashionable hotels, sits in the lobby, smokes and talks and writes letters on the hotel stationery. Speaks of his "famous friends," Rockefeller, et al.

6. (Complete scale.) Man, age 84; M. A. 11-7; I. Q. 73. Born in Bavaria. Finished German Volksschule. Speaks good English. Spends all his time painting portraits and scenes in oil. Witty, though humor is somewhat forced. Tries to be funny in all his responses, and attempts to conceal lack of ability by a pun or joke. In trying to repeat the 60 words, kept interrupting with the remark, "Ach Gott, ain't that so foolish. That's nothing to say; it's foolish."

7. (Complete scale.) Woman, married, age 59; M. A. 11-9; I. Q. 74. Bohemian; speaks English fluently. Attended school about five years. Former occupation—housework, nurse, cook in restaurant. Diseases—Syphilis—locomotor ataxia (walks with cruches; can hardly stand); cutaneous

aphasias, skin hypersensitive in spots; glandular affection, over-secretion of saliva.

8. (Abbreviated scale.) Woman, age 81; M. A. 12-0; I. Q. 75. Attended school only about a year. Says there are not many with whom she would exchange places. In Home 4½ years. Former profession—millinery. Had a stroke of apoplexy in 1893 which left her temporarily unconscious, blind and paralyzed. It affected her speech for six months. Vision now defective. Suffers some from neuralgia. Never read a book through in her life. Spends time now reading the Bible and newspapers. Absent minded. She asked three times to have the purpose of the test explained to her. Talks so incessantly it was difficult to keep her to the task at all. Consequently the testing was very slow.

9. (Abbreviated scale.) Man, age 85; M. A. 12-2; I. Q. 76. Went to school until 15 years old. Blacksmith. In Home 4 years. Reads a good deal, chiefly the Bible and religious papers. Does some gardening. Wants to get married.

10. (Complete scale.) Woman, age 65; M. A. 13-2; I. Q. 82.3. Vocabulary score 76. Completed 8th grade. In Home 4 years. Profession nursing. Very disagreeable in the Home. Constantly quarreling with inmates, and making it unpleasant generally for the others in the same cottage. Abnormally neat and trim. Mania for spotlessness. Spends whole day cleaning the little cottage. Gets mad with rage and goes into hysterics if a towel is left out of place.

11. (Abbreviated scale.) Woman, age 81; M. A. 13-4; I. Q. 83. Vocabulary score 68. Attended a seminary two years after completing elementary school. Had great ambition to be a teacher. In Home 5 years. Totally blind from an unsuccessful operation for cataract. For this reason it was necessary to select from each group such tests as did not require vision. It was possible to choose four or five from each group which she could take. Sweet spirited. Very religious. Has different ones come in and read to her. Has a keen wit and sense of humor, and is youthful in spirit.

12. (Abbreviated scale.) Man, age 76; M. A. 13-9.5; I. Q. 86.5. Vocabulary score 86. Attended school about five years, until 16 years old. Has had typhoid fever, rheumatism, hernia, and heart trouble. Somewhat deaf, and vision poor. Large tumor on back of neck. Very restless. Would get up and roam around during the test. Speaks with quick, nervous accent. Also very quick in decision and response. Says he

remembers events that happened 50 or 60 years ago better than those of recent occurrence. Says he can tell where he spent every Christmas in his life.

13. (Complete scale.) Woman, age 59; M. A. 14-2; I. Q. 89. Vocabulary score 81. Attended school until 17 years of age. Went to about the 7th grade. Has had typhoid fever, influenza, pleurisy, and tuberculosis. Reads Scribners, Century, and Harpers Magazines, and some of Dicken's novels. Has a large vocabulary (score 81). Totally lacking in self-confidence; is perfectly dependent upon her mother, who is also in the Home. The superintendent of the institution says she depends upon her mother as a child. Her mother says that the daughter always asks her what dress to put on, etc. Neurasthenic. Formerly suffered from hysteria. Very dirty in habits of life. Sometimes eats from garbage can. Curt, impudent, and unappreciative towards authorities. The mother (No. 19) recognizes her daughter's abnormality, and says she is mentally affected. The psychiatrist in charge of the institution says she is abnormal.

14. (Abbreviated scale.) Man, age 93; M. A. 14-8; I. Q. 92. Vocabulary score 83. German; speaks English fluently. Education—7th or 8th grade. Former school teacher. Reads current papers and writes poetry for pastime. Slips off from the Home and begs in town, though the Home amply provides for every need, including tobacco, and materials for ladies' fancy work. Has been in jail once for begging and threatened several times. When finally told that he would be put out of the Home if he persisted in begging, he said, "Well, I guess I'll have to quit that." Slightly deaf, eczema, digestive trouble. Abnormal sex tendencies.

15. (Complete scale.) Man, age 87; M. A. 15-0; I. Q. 94. Vocabulary score 82. Went to about 7th grade in school. Ranchman. Good character. Sense of humor. An intelligent looking and intelligent acting man, cordial, hearty, and youthful in manner. Spoken well of by director. Took great interest in the test.

16. (Abbreviated scale.) Man, age 76; M. A. 15-3.5; I. Q. 95.5. Vocabulary score 74. Completed elementary school. Contractor. In Home 5 years and 3 months. Spends time gardening. Has a beautiful yard and extensive garden, and takes great pride in it. Very successful in raising all season strawberries. Makes a good deal by selling these to tourists and visitors. Good natured, witty, and attractive personality. Likes to talk, and talks intelligently. Reads a great deal.

Is liked by inmates and authorities. Eyes defective, and somewhat deaf. Had pleurisy at age of 23.

17. (Complete scale for more difficult tests, but abbreviated scale for easier ones.) Woman, married, age 69; M. A. 15-11; I. Q. 99.4. Vocabulary score 88. Has suffered recently from a severe fall from which she was confined in bed 18 months; also from rheumatism, a tumor operation, and nervous trouble. Formerly hired girl. Husband was a farmer and miner. In Home 4.5 years. Very religious. Garrulous. Reads religious papers and current magazines. Likes the National Geographic Magazine and the Literary Digest especially. Interested in the Bible and in history. Says she always had a vivid imagination, and for that reason did not read novels. Reads to the blind lady, No. 11.

18. (Complete scale.) Woman, age 80; M. A. 15-11; I. Q. 99.4. Vocabulary score 80. Born in Boston. Reached the third year of high school. Former nurse. Complains of the institution a good deal. Was put out of another institution for such complaints. Took great interest in the purpose and results of the tests. Continued asking questions about them nearly every day. Encouraged others to take the test. Feels her superiority in the Home. Is considered peculiar by the inmates. Enjoyed the examination, but laughed almost hysterically during most of the test. Is a Christian Scientist, and reads the literature a good deal. Suffers from growing cataract.

19. (Abbreviated scale.) Woman, married. Mother of No. 13. Age 79; M. A. 16-3; I. Q. 102. Vocabulary score 74. Attended school until 15 or 16 years old. Says she led her classes. Was married twice; husbands both farmers. Previous occupation, nurse, cook in restaurant; ran a laundry. Occupies herself now with reading, sewing, knitting, and fancy work. Growing cataract. Bears tales to the authorities and to the town people. Keeps cottage in a very dirty, filthy condition. Treats her 59 year old daughter (No. 13) as a spoiled child. Was very apologetic for the daughter; explained her condition and the "cause," and asked permission to observe her daughter take the test.

20. (Complete scale for more difficult tests, but abbreviated scale for the easier ones.) Woman, age 68; M. A. 17-6; I. Q. 109. Vocabulary score 80. In Home 4.5 years. Completed district school, and then attended a ladies' seminary from the ages of 15 to 18. Studied music as a child in Chicago, then in the ladies' seminary, then one year in the New Eng-

land Conservatory of Music in Boston, and under private teachers in Chicago when about 22 years old. Profession, music teacher for 22 years. Came to the West for her health 35 years ago. Her health broke down several times while in the West. She had to support herself. "Life was a little too hard for my strength," she said. Now suffers from neuritis. Devotes her time to work on the place, and to sewing and china painting and mending. Was for a time very successful in china mending, by a process which she invented. Has a cottage and kiln in town, and still makes some money from that occupation.

The average age of the group is 75 years (74.95). Their average mental age is 13-3 (13-2.98), and their average I. Q. is 83 (82.92). The average age of the men alone is 80.3 years. The average age of the women is 69.6 years. The average mental age of the men is 12-5.8. The average mental age of the women is 14 years and 1/10 month. The average I. Q. of the men is 78.21, as compared with 87.65 of the women. Thus the women have an average mental age of 1 year and 6.3 months more than the men, and an I. Q. of 9.44 points higher than the men, but the men are on an average 10.7 years older than the women.

In order to ascertain whether this difference in favor of the women is due to the greater age of the men, or to sex differences, or to selective factors, several comparisons were made. It was impossible to get any information on the effect of age alone, regardless of sex, by comparing the five or six oldest in the group with the five or six youngest, because the five oldest were all men and the five youngest all women.

Accordingly, in order to eliminate the age factor and get at the difference in sex alone, the five oldest women were compared with five men of approximately the same ages. The five men chosen were numbers 3, 4, 5, 12 and 16. The average age of the five men is 77.2 years; that of the five oldest women is 78 years. The average mental age of the men is 12-7 (12-6.96). The average mental age of the women is 14-8.16. The average I. Q. of the five men is 78.8. The average I. Q. of the five oldest women is 91.8. So evidently the difference in intelligence between the women and the men is not due to the greater age of the men. As a check, the five oldest women were compared with the five youngest men.

The average age of the five oldest women is 78 years; that of the five youngest men is 74.4. The average mental age of the five oldest women is 14-8.16; that of the five youngest men is 12-1.4. The average I. Q. of the five oldest women is 91.8; that of the five youngest men is 75.8. The difference in the intelligence of the men and women is greater in this case than when those of equal ages were compared.

Now the five oldest men may be compared with the five youngest men, and the five oldest women with the five youngest women in order to ascertain furthermore the influence of age on the mentality of these groups. The average age of the five oldest men is 86.2 years; that of the five youngest men is 74.4. The average mental age of the five oldest men is 12-10.32. The average mental age of the five youngest men is 12-1.4. The average I. Q. of the five oldest men is 80.6; that of the five youngest men is 75.8.

The average age of the five oldest women is 78 years; that of the five youngest women is 61.2 years. The average mental age of the five oldest women is 14-8.16; that of the five youngest women is 13.4 (13-3.96). The average I. Q. of the five oldest women is 91.8; that of the five youngest women is 83.46.

Finally the ten oldest of the entire group, without regard to sex, were compared with the ten youngest of the entire group. The average age of the ten oldest of the group is 82.1 years. The average age of the ten youngest of the entire group is 67.8 years. The older group was composed of 5 men and 5 women; the younger group also of 5 men and 5 women, the five oldest in each case being men. The mental age of the ten oldest is 13-9.26; that of the ten youngest is 12-8.7. The average I. Q. of the ten oldest is 86.21. The average I. Q. of the ten youngest is 79.63. The same fact is demonstrated again, notwithstanding the fact that the one who made the highest score (109) is only 68 years old, and has been in the younger groups throughout.

Instead of the expected deterioration in the mentality of the older ones, due to senescence, it appears that for some reason the older ones are more intelligent than the younger ones. There is evidently some selective factor operative among inmates of such institutions which is stronger than the influence of age alone between the ages of 55 and 93. This fact appears constantly in all of the groups arranged for age comparisons: the ten oldest compared with the ten youngest, the five oldest women compared with the five youngest men, the five oldest men compared with the five youngest men, and the five oldest women compared with the five youngest women. Only in a comparison of the five oldest men with the five youngest women is the advantage on the side of the younger group, and this is due to the much higher intelligence of the women than of the men tested.

This greater intelligence of the older inmates of the institution is probably due to the fact that the dull ones are unable to support themselves so long as the bright ones, and hence become dependent earlier. The more intelligent ones are probably able to earn a livelihood and provide and care for themselves longer, and consequently keep out of the Home longer, than the less intelligent ones. The more able ones finally come to the institution, no doubt, chiefly on account of old age (i. e. physical decrepitude), while the younger ones may come because of their inability, at a relatively young age, to make a living. It may be also that the more intelligent ones dread accepting charity more than the less intelligent, and consequently try more eagerly to avoid getting into a Home for the Aged, or to postpone it as long as possible, while the less intelligent probably care less, and consequently become a public burden at an earlier age. The fact that the older ones have survived sickness and death longer may also be an indication of greater mental and physical vigor.

Notwithstanding these facts, however, the women, though considerably younger than the men, are also much more intelligent. This might be a sex difference; it might be due to a slower deterioration of the mental processes of women than of men at senescence, but it is more likely due to the fact that women are paid less for their services and consequently find it more difficult to make a living than men, since relatively more is demanded of the woman worker for the same wages. There is probably also less opportunity for old women to earn a living than for old men. Consequently, though relatively intelligent, they come to the Home early because of inability to support themselves. Most of the women (6 out of 10), too, are single women, and had to maintain themselves. Some are there because of illness, and others because of the early death of their husbands. In only one case was the husband still living with the woman. So some were likely suddenly thrown upon their own responsibility for the first time by the death of their husbands, and were unable, because of inexperience, to make a living. For these reasons, and because of the greater number of old women than old men existing, there are more women in such institutions than men.

In order to determine the influence of education upon the ability of the inmates of this institution to earn a living and to remain self-supporting until well advanced in years, comparisons were made on the basis of school training and age. It was found that the ten older ones of the entire group attended school on an average about 7.9 years, and the ten younger ones 7.33 years. The men attended school 7.8 years and the women 7.33 years. The five older men attended school 7.8 years, and the five younger men also 7.8 years. The five older women attended school an average of 7.5 years; the five younger women 7.2 years. Since the older group attended school as long as or longer than the younger group in all cases except where men and women are compared separately, this seems to indicate a greater love of school work by the older or more intelligent ones. The difference in the school training of the two groups, however, is so small as not to warrant the inference that those who have attended school longer are more successful in life because of their slightly superior education.

In order to ascertain what mental processes offer special difficulty to the old people, and which ones are easy, the following table is given, showing the number of subjects who passed each test correctly, and the number and percentage of those who failed on each test. In only two cases was it necessary to begin lower on the scale than age VIII. In group VII the 3rd test (repeating 5 digits) was passed by one woman and missed by one man.

TABLE I

	Tests	No. of Cases Answering Test Correctly	No. of Cases Failing on Test	Per cent Failing on Test
	Age VIII			
1.	Ball and field	2	3	60
2.	Counts 20–0	12	0	0
3.	Comprehension, 3rd. degree	12	0	0
4.	Gives similarities (2 things)	12	0	0
5.	Definitions superior to use	6	0	0
6.	Vocabulary, 20 words	11	0	0
	Totals	55	3	5.17
	Age IX			
1.	Date	12	2	14.28
2.	Arranging 5 weights	5	9	64.3
3.	Makes change	7	0	0.
4.	Repeats 4 digits backwards	9	6	40.
5.	Three words in sentence	11	3	21.43
6.	Rhymes	4	3	42.86
Al.	1. Months	1	0	0
Al.	2. Stamps	1	0	0
	Totals	50	23	31.5

From this it appears that the mental processes in which old people are most often deficient are, in the order named: immediate auditory memory for digits and for words, visual memory or perception, analysis and synthesis, inventiveness and ingenuity, motor co-ordination, imagination or perception, using and manipulating numbers, arithmetical reasoning, visual imagery, comprehension, logical memory, knowledge, abstracting common element—generalization and interpretation, quantitative comparison (kinaesthetic sensory discrimination), qualitative comparison, defining—association and analysis.

Of all the more important functions, the old people are best by far in vocabulary. Eight out of the 20 subjects passed the superior adult vocabulary test of 75 words, all but one of these defining 80 words or more. The best two in the vocabulary test made scores of 88 and 86 respectively. These did not, however, make the highest scores on the whole scale. Their intelligence quotients were 99.4 and 86.5 respectively (case numbers 17 and 12). Three others passed the average adult vocabulary test of 65 words—intelligence quotients 102, 95.5, and 83.3 (case numbers 19, 16, and 11), the latter a blind lady, age 81.

In all of the 25 more difficult tests listed in Table II the women excel, with the exception of the induction test (men 50% failures, women 100% failures), but this was given to only 2 men and 1 woman, and consequently the figures have little significance; and the ball and field test, inferior plan (men 50% failures, women 100% failures; 4 men and 1 woman).

The tests in which the women excelled very greatly were, in the order of the difference in the results of the men and the women:

	Test	Per cent Failures by Men	Per cent Failures by Women	Difference
1.	Arrangement of 5 weights	87.5	33.33	54.27
2.	3 words in sentence (7 men and 7 women)	43.	00.	43.
3.	Repeating thought of passage heard	100.	57.1	42.9
4.	Rhymes	66.66	25.	41.66
5.	Reversing hands of clock	100.	60.	40.
6.	Problem of enclosed boxes	85.7	50.	35.7
7.	Vocabulary, 65 words	55.5	30.	25.5
8.	Dissected sentences	57.1	33.33	23.77

CONCLUSIONS

1. The Homes for the Aged are selective agencies. Some of the most important factors of selection operative among the inmates are probably: improvidence, indolence, lack of foresight, ambition, energy, and persistence, lack of force and strength of character; ill health or misfortune, bad habits and, among those very low in intelligence, their dullness or lack of capacity. Also whether or not they have children or relatives who can supoprt them, is doubtless an external factor, independent of traits of character or intelligence.

2. The old people who are cared for by even the most selected Homes for the Aged are dull, average I. Q. in this one about 83. It is very probable that their dullness is due, to a great extent, to a decline in the mental processes at senescence.

3. It is not possible by studies of the inmates of institutions or Homes of this kind to draw conclusions with regard to the average intelligence of the aged or the decline of the mental processes at senescence.

4. The women in this Home average about 11 years younger than the men.

5. The women in such institutions are, for some reason, considerably above the men in intelligence. This may be due partly to a slower deterioration of the mental processes of women in old age. It is very probably due in part to the fact that women receive less pay for the same quality of work than men, and consequently require greater intelligence than men in order to earn the income necessary for self-support.

6. The older half of the entire group, as well as the older half of the men alone and of the women alone, are more intelligent than the younger ones. This is likely due to the fact that the more intelligent ones can probably remain self-supporting longer than the less intelligent ones, and for that reason come to the institution later than these. Probably also they endeavor more to avoid accepting public charity. They are also more highly selected mentally and physically.

7. The younger, or less intelligent half of the women, are more intelligent than the older, or more intelligent half of the men.

8. The men have, on an average, somewhat more school training than the women. The older half have generally had more school training than the younger half. This would indi-

cate that the more intelligent ones probably enjoyed school work more.

9. No deterioration in mentality with advancing age could be found, because of selective factors, chief of which is probably the one mentioned under caption 6.

10. The mental deficiency is greatest in immediate auditory memory for digits and for words, and in immediate visual memory or perception, analysis and synthesis, inventiveness and ingenuity, imagination, manipulation of numbers, and arithmetical reasoning.

11. The women were superior to the men in all the more difficult tests except two: the induction test, and the ball and field test (inferior plan).

RACIAL DIFFERENCES IN MENTAL FATIGUE[1]

By THOS. R. GARTH, University of Texas

During the year of 1913-14 I made a study of mental fatigue during the continuous exercise of a single function which was the addition of columns of one-place numbers upon 711 children of the schools of several school systems of the State of Virginia. About half of these were from the Third and Fourth Grades and the others were from the Seventh and Eighth Grades. From these I derived group curves. The description of the experiment and the interpretation of the results are to be found in No. 41 of the *Archives of Psychology*, Columbia University, 1918.

Later I thought it might be worth my while to give the same tests to negro children of the same grades and likewise to Indians. The negroes were tested in the spring of 1917 in the Baker School of Richmond, Virginia. In the spring of 1919 I had the opportunity to give the same tests, by the courtesy of Mr. O. S. Lipps, Supervisor of the United States Indian School at Chilocco, Oklahoma, to two similar groups of Indians.

The experiment consisted of adding the Thorndike Addition sheets continuously over twenty-eight minutes of time for the younger groups and forty-two minutes for the older ones. I should say that the only break in the continuity was the time taken in going from one sheet to the next one and this break was not more than half a minute in duration. The subjects were allowed two minutes to a sheet. The motivation consisted in the experimenter's saying at the beginning of the performance that he wished to find out the best "adder" in the class.

The Indians.

In the Indian groups there were boys and girls of thirty-four different tribes whose degree of blood ranged from ⅛ Indian to full blood. The Cherokees numbered 39; Creeks, 21; Cheyennes, 14; Choctaws, 12; Chickasaws, 12; and Caddos, 13. In the various other tribes the numbers are too small to be given here. As to degree of Indian blood there were 91 full bloods, 8 of ⅞ degree, 28 of ¾ degree, 4 of 9/16, 47 halfbreeds, 10 of ¼, and 1 of ⅛. It will be seen that nearly half of the Indians were full bloods. One hundred twenty-seven were ¼ Indians and more. These figures, I understand are taken by the United States government and I believe are fairly reliable.

---

[1] Read before the Southern Society for Philosophy and Psychology at the New Orleans meeting, April 23, 1920.

The Negroes.

No attempt was made to ascertain the purity of racial blood of the negroes. There were many mulattoes among them, but I believe that the difference in fatigue shown here is significant as pointing to racial difference in continuous performance of a mental act.

In the Third and Fourth grade groups there were 368 whites, 92 Indians, and 99 negroes. The average age of the whites was 9.9 years, Indians 14.8 years and negroes 11.5 years.

In the Seventh and Eighth Grades there were 343 whites, 98 Indians, and 34 negroes.

The average ages of the above racial groups are: whites, 14.05 years; Indians, 17.9 years, and negroes 14.3 years.

We should also state that there were five Indians of the younger group in the fifth grade and that all the negroes are from the seventh grade in the older negro group. It was impossible to obtain enough Indians at the Chilocco school to make even hundreds in both groups. However, we do not believe that these discrepancies affect the results to any appreciable degree.

Since the total performances of the racial groups are a matter of interest they are given herewith, though they are not presented as a part of the essential data from which to draw conclusion at present because the purpose of the experiment was to study not total performance, either attempted or accurate, but the fatigue of the groups comparatively. The total performance was—

For Third and Fourth Grade Groups

Attempts				Accurates			
Whites	Ave.	46.6	columns	Whites	Ave.	26.6	columns
Indians	"	49.5	"	Indians	"	27.0	"
Negroes	"	51.6	"	Negroes	"	32.2	"

For Seventh and Eighth Grade Groups.

Attempts				Accurates			
Whites	Ave.	123.60	columns	Whites	Ave.	84.80	columns
Indians	"	116.35	"	Indians	"	68.20	"
Negroes	"	117.82	"	Negroes	"	84.38	"

Average of the Individual Curves.

In the dissertation referred to above the writer states on page 9: "The object in a study of this kind is to give every individual, whether his work is slow or fast, the same influence upon the final average. It is not fair to the poor or weak individual to put his small effort on a basis of absolute terms along with the effort of the better workers similarly expressed. The original data (absolute measures)—fail to represent truly the facts which we wish to bring out because, as intimated above, the rapid workers with their larger measures of performance entirely swamp the changes that may occur in the work of the slower ones. But if everyone's work for the successive periods is reduced to percentages of his total, the slow count the same as do the fast workers in determining the general work curve. For this reason we have derived percentage curves which represent more truly—(than the absolute measure curves) the central tendency of the whole group with regard to fatigue."

For the reason given above it was thought best to reduce the performance of each subject at each two minutes of his working to a percentage of his total performance. This curve we have called the Individual Curve. An average of these individual curves was obtained for both, attempts and accurates, for each group. Only the Averages of the Individual Curves are shown here. This seems right to us since we are not just now studying the absolute performance of the groups but the fatigue of the racial groups.

We give here Tables I, II, III, and IV, which show the data for the original unsmoothed curves.

TABLE I

DATA FOR THE CURVES REPRESENTING THE AVERAGE OF THE INDIVIDUAL CURVES—THIRD AND FOURTH GRADES

Columns attempted in each period of two minutes, expressed as a per cent. of the total number attempted.

	368 Whites			92 Indians			99 Negroes		
Period	Ave. Per C	A. D.	P. E.	Ave. Per C	A. D.	P. E.	Ave. Per C	A. D.	P. E.
1	6.97	1.94	.09	7.49	1.38	.11	7.69	2.60	.25
2	7.24	1.52	.08	6.93	1.27	.10	7.59	1.50	.13
3	7.30	1.37	.06	7.06	.62	.06	7.67	1.34	.11
4	7.24	1.40	.06	7.07	1.26	.11	7.82	1.35	.12
5	6.97	1.32	.06	6.70	.78	.07	6.99	1.30	.11
6	7.20	1.32	.06	7.41	1.07	.09	7.27	.05	.09
7	6.97	1.35	.06	6.61	1.62	.14	6.34	1.24	.11
8	7.29	1.43	.06	7.07	1.38	.09	6.66	1.32	.12
9	7.34	1.39	.06	6.48	1.27	.10	6.65	1.36	.12
10	7.28	1.44	.06	7.52	1.68	.14	7.03	1.23	.11
11	7.27	1.34	.06	7.66	2.65	.18	8.18	2.22	.19
12	6.78	1.47	.07	7.22	1.48	.12	6.93	1.23	.11
13	7.25	1.31	.05	7.55	2.76	.31	6.72	1.49	.12
14	6.90	1.62	.07	7.22	1.87	16	6.46	1.19	.11
	100.00			100.00			100.00		

TABLE II

Data for the Curves Representing the Average of the Individual Curves—Third and Fourth Grades

Columns accurate for each period of two minutes, in per cent. of total columns accurate.

	368 Whites			92 Indians			99 Negroes		
Period	Ave. Per C.	A. D.	P. E.	Ave. Per C.	A. D.	P. E.	Ave. Per C.	A. D.	P. E.
1	6.13	4.53	.20	7.65	3.54	.32	8.80	3.38	.27
2	8.34	4.87	.22	6.17	3.00	.27	7.93	3.47	.28
3	8.16	4.34	.21	7.95	3.42	.28	8.93	3.52	.29
4	7.53	4.13	.13	3.42	4.65	.40	7.70	3.04	.27
5	7.27	4.20	.19	8.23	5.14	.45	6.62	3.48	.28
6	7.76	4.10	.18	7.26	2.73	.22	6.97	3.12	.27
7	6.94	4.01	.18	5.80	4.58	.40	6.95	2.97	.25
8	7.24	4.05	.18	7.54	3.12	.27	6.98	3.20	.28
9	7.62	4.57	.21	5.51	3.92	.30	6.57	2.78	.20
10	7.16	4.31	.19	7.92	4.17	.36	6.67	3.30	.29
11	7.06	4.27	.19	7.87	4.12	.36	7.19	4.32	.38
12	6.10	4.07	.18	6.93	3.81	.31	6.87	3.12	.27
13	6.89	4.30	.19	7.24	3.34	.29	6.27	3.03	.27
14	5.80	3.75	.17	6.15	4.63	.41	5.55	3.40	.29
	100.00			100.00			100.00		

TABLE III

Data for the Curves Representing the Average of the Individual Curves—Seventh and Eighth Grades

Columns attempted for each period of two minutes in per cent. of total columns attempted.

	343 Whites			98 Indians			34 Negroes		
Period	Ave. Per C.	A. D.	P. E.	Ave. Per C.	A. D.	P. E.	Ave. Per C.	A. D.	P. E.
1	5.00	1.01	.05	5.10	.68	.06	5.91	1.32	.17
2	4.86	.79	.04	4.68	.39	.03	4.52	.87	.13
3	4.67	.77	.04	4.61	.62	.05	4.69	.88	.12
4	4.79	.69	.03	5.12	.81	.07	4.68	.86	.12
5	4.37	.75	.03	4.64	.64	.06	4.45	1.00	.15
6	4.73	.72	.03	4.61	.67	.07	4.12	.34	.05
7	5.30	1.04	.05	4.50	.71	.06	4.79	1.81	.27
8	4.63	.79	.04	4.97	.38	.04	4.75	.71	.10
9	4.51	.75	.03	4.61	.53	.05	4.97	.74	.11
10	4.86	.75	.03	4.71	.51	.05	4.79	.58	.07
11	4.80	.68	.03	4.67	.62	.05	4.84	.80	.12
12	4.43	.74	.03	4.67	.58	.06	4.69	.74	.10
13	4.42	.88	.04	4.94	.46	.04	4.54	.44	.06
14	4.68	.83	.04	4.50	.69	.07	4.80	.92	.13
15	5.58	1.11	.05	4.81	.68	.07	4.54	.94	.13
16	4.74	.62	.03	4.72	.66	.07	4.32	1.33	.21
17	4.80	.71	.03	4.77	.83	.08	5.28	1.06	.15
18	4.75	.90	.04	5.05	.42	.04	4.91	.32	.05
19	4.40	.86	.04	4.90	.59	.05	4.54	1.06	.15
20	4.84	.91	.04	4.90	.57	.05	5.24	.51	.08
21	4.76	.90	.04	4.84	1.06	.09	4.63	1.18	.16
	100.00			100.00			100.00		

TABLE IV

Data for the Curves Representing the Average of the Individual Curves—Seventh and Eighth Grades

Columns accurate for each period of two minutes, in per cent. of total columns accurate.

	343 Whites			98 Indians			34 Negroes		
Period	Ave. Per C.	A. D.	P. E.	Ave. Per C.	A.D.	P. E.	Ave. Per C.	A. D.	P. E.
1	5.00	1.74	.08	4.93	1.90	.17	7.41	3.04	.45
2	4.85	1.51	.07	4.19	1.05	.09	4.57	1.25	.15
3	4.85	1.51	.06	5.41	1.38	.11	5.36	2.00	.30
4	5.01	1.37	.06	5.65	1.79	.15	4.05	.70	.10
5	4.50	1.37	.06	5.20	1.94	.17	4.67	2.67	.30
6	5.03	1.33	.07	5.13	1.74	.14	4.28	1.10	.15
7	4.70	1.46	.06	4.39	1.68	.13	4.43	2.35	.31
8	4.83	1.37	.06	5.16	1.82	.16	5.39	1.61	.18
9	4.73	1.30	.06	4.03	1.78	.15	5.97	2.76	.32
10	5.15	1.48	.07	5.48	1.76	.14	4.36	.29	.05
11	5.10	1.28	.07	4.58	1.89	.16	4.25	1.42	.21
12	4.34	1.45	.06	4.39	1.93	.17	4.67	1.94	.28
13	5.03	1.39	.07	5.07	2.63	.22	4.39	1.48	.22
14	4.37	1.51	.07	4.05	1.39	.11	4.28	1.14	.15
15	4.90	1.48	.07	4.29	1.48	.12	4.94	2.18	.30
16	4.74	1.55	.06	3.51	1.58	.13	5.10	1.80	.18
17	4.84	1.42	.06	5.32	1.09	.09	5.28	1.18	.16
18	4.73	1.23	.07	5.02	2.00	.18	4.28	.87	.12
19	4.40	1.35	.08	4.89	1.53	.13	4.05	2.28	.31
20	4.60	1.53	.07	4.72	1.60	.14	4.22	2.10	.30
21	4.30	1.69	.06	4.59	1.93	.17	4.05	1.92	.29

The Original Curves

The curves which we show here are not the original curves showing the Average of the Individual Curves but are the smoothed-out curves obtained by averaging the per cents for every two periods taken in succession. In the curves we find these features:

Third and Fourth Grades—Attempts Curves.

The whites show a warming up, fluctuations, reach their maximum in the ninth period or by the 18th minute; the Indians instead of exhibiting warming up show initial spurt, a falling away and recovery in the 10th period reaching their maximum in the 13th period; the negroes like the Indians take an initial spurt and maintain a fairly good rate until the 6th period when they fall away, recover and then fall away again experiencing their maximum at the 11th period.

Third and Fourth Grades—Accurates Curves

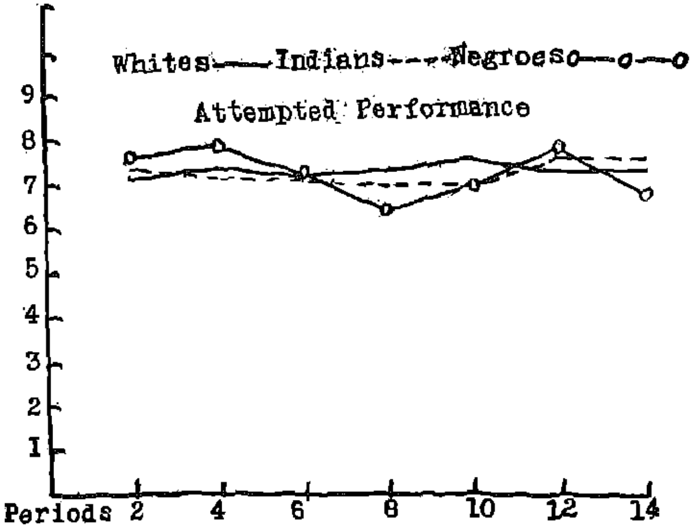


FIG. 1.—Showing averages of the individual curves.

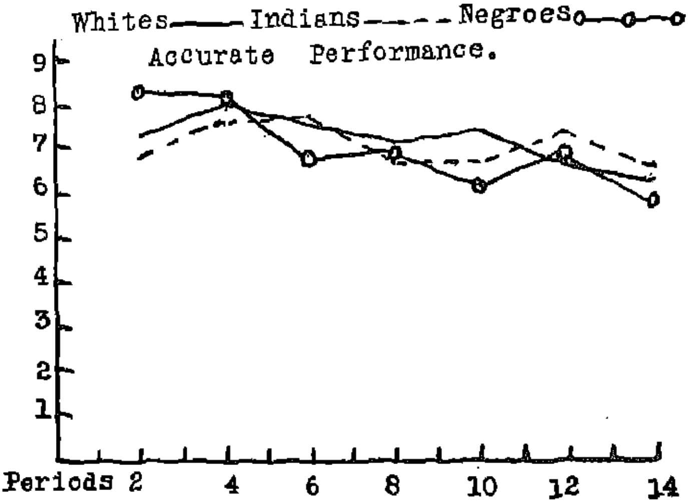


FIG. 2.—Showing average of the individual curves.

Here again the whites show warming up reaching their maximum in the 2nd period and after fluctuations fall away at the end; the Indians show as before initial spurt with maximum at the 5th period and after fluctuations fall away at end; the negroes show initial spurt, maximum at 3rd period and with fluctuations fall away at end.

Seventh and Eighth Grades—Attempts Curves.

The whites here show initial spurt reaching their maximum in the 15th period and falling away slightly at the end; the Indians show initial spurt, maximum at 4th period, fluctuate some and fall away slightly, then recover toward the end of the practice; the negroes exhibit a more decided initial spurt than was shown by the other races, the maximum occurs in the first period, fluctuations occur, a slight falling away is seen from which recovery is made at the last period but one which is followed by a falling off.

Seventh and Eighth Grades—Accurate Curves.

The whites again show initial spurt, experience their maximum at 10th period fluctuate and fall away at the end; the Indians likewise show initial spurt, reach maximum at 4th period, after fluctuations fall away slightly at end; negroes, as in attempts category, exhibit initial spurt and maximum in the first period, fluctuate and experience decided fatigue at end.

We are unable to give a satisfactory reason why the whites should reach their maximum first, then the negroes and after them the Indians in both attempts and accurates in the lower grades, and why the order for the upper grades should be for maxima, negroes, Indians and then whites. If an early occurring maximum indicates early fatigue, then the whites showed evidence of fatigue first and the negroes next in the younger groups. Should this be true the whites must have recovered much for they were able to resist the onset of fatigue better at the close than the negroes as will be shown later. As for the upper grades the order is not what we would expect as the Indians reach their maximum before the whites. Recovery on the part of the Red group might be the explanation.

Fatigue.

Since fatigue is indicated by a falling away from previous performance we have taken here the performance in the first three minutes as the standard and compared that with same number of minutes at the end.

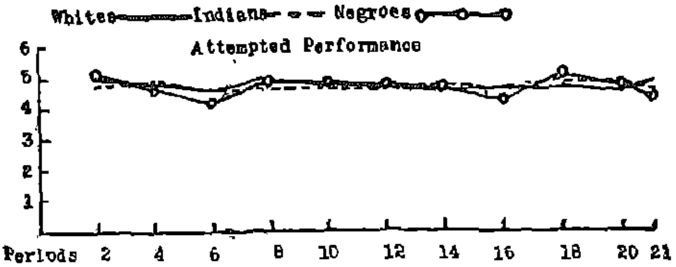


FIG. 3.—Showing average of the individual curves.

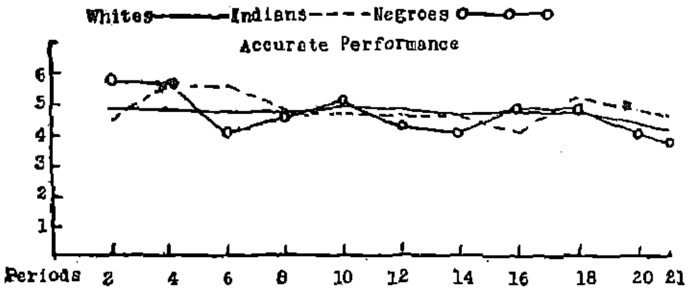


FIG. 4.—Showing average of the individual curves.

These numbers show that for the younger groups

In the Attempts		In the Accurates	
Indians gained	3.25%	Indians lost	8.4%
Whites lost	3.00%	Whites lost	17.0%
Negroes lost	12.10%	Negroes lost	27.2%

In both categories, attempts and accurates, when we arranged the races in the order of least to greatest fatigue we have Indians, whites, and negroes.

For the older groups the numbers are,

Attempts:		Accurates:	
Indians gained	2.10%	Indians lost	1.30%
Whites lost	4.00%	Whites lost	10.20%
Negroes lost	4.70%	Negroes lost	30.30%

While the evidence is not so clear as respects whites and blacks in the matter of attempted performance because the latter are only .7% worse than the former, the order seems to be the same as in the younger group, i. e., Indians, whites and negroes in accurates, and the Indians excel both races in resisting fatigue in attempted performance.

The Age of the Indian Subjects.

In giving the experiment and while working up the material the superior age of the Indians seemed to me to make it unfair to place them in comparison with the whites and negroes. There are two ways in which to meet this difficulty—(1) ascertain the effect of age on tendency to fatigue, (2) take the subjects white and red, age for age, and compare them to see if the age group of one race fatigued more than the same age group of the other race.

As respects (1) the pertinent question is, What effect does age have on resistance to fatigue in the Indian group? To answer this a correlation between the age and the resistance to fatigue for that age-group was obtained. In the Third and Fourth Grades, the correlation for attempts is a negative one of —.26, for accurates it is +.03. For the Seventh and Eighth Grades the correlation for attempts is a negative one of —.20, for accurates it is —.31. These correlations are either all negative or so small that they are of very slight significance and could be ignored.

In order to make the comparison suggested by number (2) *it would be necessary to have larger groups of Indians than* we could obtain for the age group comparison. However, we give a few comparisons.

In the lower grades 10 whites and 8 Indians of 13 years show an age-group loss of 35% for the former and a gain of 8% for the latter in attempts, for accurates it is respectively loss of 37% and loss of 5%. 10 whites and 12 Indians of 14 years show an age-group gain of 10% for the former and 14% for the latter in attempts. In accurates it is for this age for the races respectively loss of 22% and gain of 1%.

In the upper grades Attempts and Accurates 85, 77, 42 and 4 whites and 7, 18, 13 and 25 Indians belonging to the age-groups of 14, 15, 16 and 17 years respectively showed these group changes for attempts—

14 yrs. Whites gain of 1%	Indians gain of 7%
15 yrs. Whites loss of 3%	Indians gain of 8%
16 yrs. Whites loss of 4%	Indians loss of 9%
17 yrs. Whites loss of 9%	Indians gain of 5%

These changes are shown for accurates:

14 yrs. Whites loss of 7%	Indians gain of 6%
15 yrs. Whites loss of 11%	Indians gain of 2%
16 yrs. Whites loss of 22%	Indians loss of 21%
17 yrs. Whites loss of 12%	Indians loss of 4%

In the above comparisons the results are favorable to the Indians excepting in the case of the 16 year group in the attempts category where the whites lost 4% and the Indians 9%, and in the accurates category where the advantage of the Indians over the whites is too slight to be noticed particularly, the numbers are 22% and 21% respectively.

The conclusion to be drawn from this experiment, then, is that in a test to ascertain the difference between whites, Indians and negroes of equal educational opportunity in the matter of mental fatigue so-called, the Indians of both the young and old groups tend to excel the whites and negroes in this respect both in the curve representing fatigue in attempts and accurates performance, likewise the whites excel the negroes in both these categories.

The above conclusion is based on the averages of the groups as tendencies. It is very likely that there is some overlapping to be found in the fatigue data. Since this has not yet been measured we hesitate to say that we have here racial differences. The probability is that they are not to be found in these data even as they are not revealed in many other investigations. However, it is interesting to note that these Indians rate high on a scale of resistance to fatigue and that the negroes rate low, while the whites occupy a sort of middle ground if the average alone may be taken as an index.

A NEW APPLICATION OF PSYCHOLOGY TO INDUSTRY

By Henry C. Link, Ph. D.

The following experiment, conducted in a sporting goods factory where the use of psychological tests for employment purposes has been extensively developed, indicates an entirely new way in which tests can be applied.

An inspection shop, where some 80 girls were employed as bullet inspectors, was having difficulty with the current piece-work rate of pay. The foreman of the shop claimed that the rate was too low and that he was finding it impossible to retain his inspectors because of their dissatisfaction with their earnings. However, the time-study group whose business it was to set piece-work rates, asserted that the rate was a just one. Moreover, they showed that it was the highest rate which had ever been paid for that kind of work. An investigation proved that both claims were apparently well founded, and yet the investigation contributed nothing in the way of an intelligent solution of the problem.

For the benefit of those who do not understand the piece-work rate or how it is set it will be necessary to give a brief description. The piece-work rate is a rate of remuneration which is governed by the quantity of work an individual performs. It is a rate which enables each worker to profit directly according to his capability. This rate, however, must first of all be determined, and this process requires a special study of the work to be done and of the comparative capabilities of those selected to do it. The men who make this study and who then set the rate are usually called time-study men from the fact that their study consists largely of timing, by means of a stop watch, the elements of the work or operation for which the rate is to be set.

For example, in setting the rate for bullet inspectors, a time-study man first analyzes the operation into its elements, perhaps as follows:

1. Picking up a handful of bullets.
2. Inspecting bullet points.
3. Reversing bullets and inspecting bases.
4. Rolling bullets out of left hand into right and inspecting lateral surface.

This constitutes a complete "cycle" in the operation of bullet inspecting. By means of a stop watch, the time-study man proceeds to time each element of this cycle, timing as many cycles as he may think necessary. He may time the work of an inspector for one or two hours or even half a day, depending on the difficulty or importance of the rate to be set. In this way he obtains both the total time it takes the inspector to perform a certain amount of work, and also the various times required by each element in the operation. He then repeats his observations on the work of several more inspectors. On the basis of the results which he obtains from these time studies, the time-study man computes what he regards a fair average rate of production, or the amount of work which a girl of average ability may fairly be expected to do. On this basis, after certain allowances of time for fatigue, personal necessities, and other possible interruptions have been made, the rate is set.

Rate setting by means of a time study is, therefore, in one sense, a psychological experiment, for it attempts to make an accurate record of a series of well-defined reaction times, and on this basis to compute the time for a series of typical reactions. However, it can readily be seen from the above description of the time-study process that the rate resulting from such a study depends almost entirely upon the subjects selected for observation. If the subjects selected are representative of the group to which they belong, the study is likely to have a valuable result. If the subjects selected are not representative—and this is frequently the case—the resulting rate will be too high or too low. For instance, a piece rate for fifty or more workers is often based upon a study of only two or three subjects. No sound statistical procedure is applied to control the great possibilities of error which may here arise. Moreover, there is an equal lack of statistical soundness in the method by which the subjects who serve as the basis of the time-study are chosen. It is quite customary for the time-study man to go to the foreman and ask him to pick out two or three representative girls, one fast worker, a slow worker, and an average worker. The foreman, who most frequently has his own axe to grind, naturally picks out those workers who will be most likely to give the result he himself desires.

In the instance mentioned at the opening of this paper, the rate set had proved unsatisfactory. However, it was impossible to determine whether the rate had been set too low; or whether the girls who were doing the work were of unusually low ability; or in fact, where the real difficulty lay. The experiment described here was conducted in an attempt to cast further light on this problem.

The eighty bullet inspectors were all employed at the same kind of work. The first object of the experiment was to discover how these girls compared with other inspectors who were doing similar work in other shops of the same company. The comparison was to be based upon a set of tests which had already been given to large numbers of inspectors in other shops. Two tests were given to each girl, the Woodworth-Wells cancellation test and the Woodworth-Wells number group checking tests. These two tests had been found very significant in previous experiments with inspectors, having shown consistently high correlations.[1] The two tests were given to each girl at her work bench and occupied about ten minutes for each subject. Forty-two of the eighty inspectors, selected at random, were tested in this way. It was taken for granted that the other thirty-eight subjects would not materially alter the averages obtained. The average performance of the forty-two bullet inspectors in these two tests was then compared with the average performance in these tests found for other groups of inspectors. The figures on which this comparison was based are given below, the results of the present experiment being put last:

Inspecting Shops	Tests Given in 1917	No. of Girls Tested	Av. Perf. in no. Group Test	Av. Perf. in Canc. Test[2]
A	April	50	161	122
A1	Nov.	66	174	136
B	June	12	175	122
C	June	22	183	160
D	May	30	187	136
E	Dec.	42	191	139

The comparison showed that the bullet inspectors were poorer than any other group of inspectors in the number group checking test (the test which had shown the most significant correlations in previous experiments), and poorer than any other group but one in the number cancellation test (the next most significant test). It was inferred from this that the dissatisfaction of the bullet inspectors was probably due to this comparative inferiority rather than to the lowness of the rate. However, the experiment did not conclude with this inference.

---

[1] The details of these experiments are described in the writer's book on *Employment Psychology*.

[2] Corrected time in seconds.

After a girl had been tested and the observer had picked up his materials, he would ask in a casual way, "How are you making out?" Of the number of girls asked, sixteen expressed themselves as not doing well or as being often discouraged, and only five expressed satisfaction. The rest were incoherent or inconclusive. This pointed to a more or less general dissatisfaction among the girls. It was obviously unfair to attribute all of this dissatisfaction to the existing rate. Further observation revealed the following contributing causes:

First, when a box of improperly inspected bullets was found, it was returned to the inspector for reinspection and had to be inspected again without pay. This was customary and expected. However, in the Bullet Inspecting Shop the boxes to be re-inspected were usually held back until 5 o'clock in the evening and only then returned to the careless inspector. In this way a girl, laboring under the impression that she was about to earn a good day's pay, might suddenly find herself, near the end of the day, confronted with a box or two of bullets which would have to be done over without pay. This spoiled her day and was quite likely to make her discouraged. It was therefore recommended that all work to be re-inspected should be returned to the inspector as soon as possible after being discovered, in order that the punishment might be as closely associated with the mistake as possible and in order to avoid ending the day with a discouraging incident. This recommendation was followed. In addition it was ordered that no bad work should be returned for reinspection after 3:00 P. M.

Secondly, it was observed that the bullets to be inspected were given to the inspectors in boxes holding as much as eighty pounds or more. It took about two hours to inspect one box, and five or six boxes constituted a day's work. When the check inspector found a few bad bullets in a box that had been inspected, the whole box must be returned for reinspection. This meant that about one-fifth of the inspector's whole day's work had been in vain and consequently she was likely to become still more discouraged. It was therefore recommended that the quantity of bullets in a box should be cut in half, and forty instead of eighty pounds given to a girl at one time. In this way, ten or twelve boxes would constitute a day's work; and if a girl worked carelessly at one time, so that a box of bullets was returned to her for reinspection, she would suffer only half as much as before. At the same time she would receive a warning which would give her an opportunity to avoid repeating her mistakes.

Thirdly, it was observed that the inspection required unusually fine eyesight. Investigation showed that 19% of all the work inspected was returned for reinspection. This was a factor to be taken into consideration in making a final adjustment of the rate.

Fourthly, the observer discovered that a practical embargo had been placed on the amount of work the inspectors were allowed to do. The limit set was seven hundred pounds. This limit was set on the supposition that a girl could not inspect *more than seven hundred pounds and do them well.* Aside from the fact that this ruling was in total contraversion of the piece-work principle, it also served as a discouragement to the very fast inspectors, who chafed under such a limitation. This limit was removed.

These observations, while not strictly psychological—that is, in the technical sense—nevertheless were directly traceable to the psychological attitude. Moreover, they were a distinct contribution to the solution of the problem.

If the tests had not been made, it might have been arbitrarily decided that the rate of pay was high enough, in which case the work of the shop might have suffered a decided detriment. Or it might have been arbitrarily decided that the rate was too low, in which case the production of the shop might not have been increased and the real root of the difficulty would have been ignored. Either course would have proved unsatisfactory to one group of claimants or the other. The results of the experiment gave a concrete and well-founded reason for maintaining the existing rate, at least for the time being or until the effects of the supplementary changes recommended could be observed. As a matter of fact, it was unnecessary to change the rate, in spite of the fact that results in the tests indicated that the bullet inspectors were not as capable as the other inspectors.

This experiment suggests certain broad potentialities of industrial psychology. The problem of setting rates is a universal industrial problem, and it is one of the most delicate and trouble-making questions with which industries have to deal. The time-study method, which is aimed chiefly at increasing production, can be distinctly supplemented by the psychological method. There is no antithesis or contradiction between the two methods for, as the above experiment clearly shows, a psychological study as well as a time study, may contribute materially to production. Moreover, in contrast with the usual procedure, an increase in satisfaction and in production may sometimes be effected without an *increase in the rate of pay.*

TENTATIVE NORMS IN THE RATIONAL LEARNING TEST[1]

By JOSEPH PETERSON, Peabody College for Teachers, Nashville, Tenn.

The present test was devised by the writer as a special study in learning, primarily. It was soon found that it is a rather good means of determining one's intellectual ability and also that it has useful diagnostic values as to certain kinds of traits, particularly degree of rational organization and of subjectivity of mind.[2] "The reaction required of the subject is to associate in a random order the numbers 1 to 10, inclusive, with the first ten letters of the alphabet. This is to be done by means of a series of guesses the range of which may be greatly limited by the use of a rational organization of the situation. Each subject completes the learning at a single practice period, varying in length inversely with the subject's ability, roughly speaking. As will be seen, the subject is forced to react to a changing situation, each response making it different to a slight degree by limiting the range of probability." The test at present can be used only as an individual test, but in this use it has some advantages over other individual tests: it is easy to give to a subject, the methods of presentation of the situation and of scoring results being objective. No printed forms other than a paragraph of instructions are necessary; the tester writes on the test sheet the necessary headings, the name of the subject, date, time of beginning, etc, while the subject reads the instructions. The time of beginning is the time at which the tester first calls out "A;" the subject is allowed as much time to read the instructions as is necessary, and should be encouraged to re-read at least once. Finally it is possible to increase the difficulty of the test to any desired point, by the addition of letters.

---

[1] Read before the Southern Society for Philosophy and Psychology in the New Orleans meeting, April 23, 1920.

[2] See Experiments in Rational Learning, *Psychol. Rev.*, 1918, 25, 443-467. A report on correlations with work in psychology will appear in the March, 1920, number of the *Journal of Educational Psychology*. A modified form of the experiment has been used by Sunne, *Psychol. Bull.*, 1919, 16, 262-267.

The norms in this report were obtained with the form of the test shown in Table I, ten letters being used, numbered as there shown. More detailed instructions as to the giving of the test are to be found in the first reference in note 2.

TABLE I

Name, L. D. Time beginning, 9:40; time ending 9:52

											Errors			
Letters...... Numbers....	*A* 9	*B* 6	*C* 2	*D* 10	*E* 8	*F* 1	*G* 5	*H* 4	*I* 7	*J* 3	Unclassified	Log. x	Per. *	Total
First series......	3 5 1 2 9	8 7 10 3 1 6	5 3 1 7 10 9x 4 2	5 7 10	8	7 5 4 1	10x 3 5	10x 3 1x 7 6x 9x 2x 7* 9x* 2x* 4	5x 3 7	3	35	10	3	48
Second series....	2 6 8 4 1 9	6	10 9x 2	9x 7 10	7 9x 8	7 5 3 9x 1	7 3 5	5x 9x 7 5x* 7* 9x* 10x 6x 9x* 6x* 4	7	9x 10x 6x 8x 2x 1x 3	33	18	5	56
Third series.....	8 6 9	2 4 8 10 1 7 8* 6	10 1 8 6x 3 7 5 8* 6x 5* 4 10* 2	10	7 5 3 1 8	1	7 3 4 2x 5	4	3 7	3	30	3	5	38
Fourth series....	9	6	2	10	8	1	5	4	7	3	0	0	0	0
Fifth series......	9	6	1 8 2	10	8	1	5	4	7	3	2	0	0	2
Sixth and seventh series correct..														
U.C. Errors.....	11	12	23	4	6	7	8	20	3	6	100			
x Errors (log.)	0	0	4	1	1	1	2	15	1	6		31		
* Errors (per.)	0	1	4	0	0	0	0	8	0	0			13	
Totals	11	13	31	5	7	8	10	43	4	12				144

Score: 12 min.; 7 repetitions; 144 total errors.

The following typewritten instructions are given to the subject to read:

> The letters A, B, C, D, E, F, G, H, I and J are numbered in a random order from 1 to 10. I call out the letters in their order and you are to guess numbers for each letter till you get the correct one, when I say "right." Then I call out the next letter; and so on. This continues till you get each number right, the first guess, twice in succession through the series from A to J. Then you are through. You must ask no questions, but are to use all the mental powers at your command in order to complete the learning as soon as possible. You will be ranked by (1) the total time you take, (2) the number of errors, or wrong guesses you make (every number you speak being a guess), and (3) the number of repetitions from A to J that you require for the learning. Re-read these instructions carefully if necessary to understand what you are to do. The meaning will be clearer as we go on with the experiment.
>
> Caution. When you get through with this experiment please do not say anything about it to others who may take it, as doing so might make it easier for them than for you.

When the subject is ready to start, the tester notes down the exact time in minutes and calls out "A." The subject usually stops after the first guess, and has to be told to go on till he gets the right number. In time the tester does not say "right" when the correct number is given, but will find it best just to call out the next letter. Every guess the subject clearly speaks out is recorded in a column just under the letter. It is common for subjects to reject certain guesses just after they have been made. This rejection should be noted by some symbol, but the number should be retained. People so often speak, and think afterwards! When the test is completed the time must be recorded immediately. Errors are then marked and counted, and the individual record will have the form shown in Table I. Logical errors are errors of guessing a number that has already been used for an earlier letter in the series, marked by an x in the record; and perseverative errors, marked in the figure by an asterisk, are errors of repeating a guess while reacting to any given letter. In indicating errors on the original record, logical errors may be checked and perseverative errors enclosed in a circle.

In addition to these quantitative records, it is well to note down, just after each test is completed, anything striking in the subject's method of going at the solution of the problem.

In this test several kinds of results are obtained, any one of which may prove to have diagnostic values, either of general intelligence as we now vaguely conceive it or of other rather general traits related in important ways to success in certain of the larger divisions of human activity. One does not, for instance, think of salesmen as being highly subjective-minded and over-conscientious. We have not attempted, therefore, to combine the scores in time, number of repetitions, and errors of the various kinds mentioned, into a single scale. It has been found, moreover, that some of these scores are far from being distributed about a mean in the form of the normal curve. This lack of so-called normality in distribution is particularly evident in the perseverative errors, and to a considerable extent also in the number of repetitions. In the former case the mode is at zero errors, and in the latter, at six repetitions, the smallest number of repetitions made by any subject. Neither time nor error data of the three other kinds shown in the table—uc., log., and total errors—approximate closely to the normal distribution. The reasons for this are to be sought partly, at least, both in the unfair sampling of college students as representative of the average adult mind (as is shown by a study of the regression lines, in the second reference indicated in note 2) and in the mental inequality of the steps of our scales. Illustrating this mental inequality of steps, we may say that the difference between a score of 36 and 86 total errors is greater than that between a score of 383 and 433, though the absolute difference in both cases is fifty.

To get a table of norms that presents all the possibly significant kinds of data, and also that does not rest its validity on an assumption of normal distribution, we have constructed a percentile table of the different kinds of scores shown in Table II. These norms are based on the records of 113 college students, mostly sophomores in psychology in the writer's classes. Eighty-one of the subjects were University of Minnesota students in the College of Science, Literature and the Arts, and were tested in the spring of 1918; and thirty-two were Peabody students, tested in the latter part of the winter quarter, 1920. The tests were given by the writer and an advanced student in the former case, and extended over a period of a week; and in the latter case by the writer and three students, all during the same day. Precautions were taken against

TABLE II

PERCENTILE NORMS IN RATIONAL LEARNING—ADULTS

(Based on 113 Cases)

Percentile	Minutes	Repetitions	Errors			
			Uc.	Log.	Persev.	Total
100	6	6	34	2	0	36
90	10.3	7.1	65	15	.8	86
80	11.6	8.8	81	22	2.0	111
75	12.3	9.5	90	26	2.7	117
70	13.0	10.2	95	30	3.4	129
65	14.3	10.8	103	34	4.5	146
60	14.9	11.3	117	41	7.3	172
55	15.5	11.8	131	43	8.5	185
50	16.1	12.4	146	51	9.4	204
45	17.7	13.0	160	55	10.2	223
40	19.5	13.5	170	60	11.0	235
35	20.5	14.3	194	67	12.5	277
30	21.8	15.4	210	82	14.8	310
25	23.4	17.1	225	95	17.9	331
20	26.5	17.8	256	102	20.7	383
10	34.7	20.7	356	148	30.9	523
0	60.0	33.0	780	335	61.0	1160

coaching, and there was no evidence whatever that any student coming to take the test knew anything of importance about it before seeing the instructions. Not all the students in the classes from which these subjects were taken were tested, but only those who could arrange favorable hours. Of the Minnesota classes—three sections—70 per cent, and of the Peabody classes only 60 per cent, were tested. There seems, however, to be no evidence of unfair sampling. Illness in the latter case prevented the testing of all the students. There is evidence that the Peabody group is composed of a more highly selected body of students than is the Minnesota group, for upon averaging the percentiles in time, repetition, and total error scores by the two groups, we found that the Minnesota median is passed by about 65 per cent of the Peabody group. The norms of the combined groups are probably roughly representative of college students. It is evident, of course, that the percentiles become less reliable as they are further removed from the median, the 50 percentile.

Correlations of results of the Rational Learning Test with scores in psychology examinations, and correlations of ranks in this test with ranks in the Otis Group Intelligence Scale and with ranks in the Army Alpha test, though yet based on an insufficient number of cases to have a high degree of reliability, indicate that the best measure of the subject's general intelligence is that which takes note of all the criteria mentioned in the instruction sheet—time, number of repetitions, and number of errors. But even then several methods may be used, each giving relatively different weight to these various factors and even to the different classes of errors described. For instance, we may find the percentile of the number of minutes, of the number of repetitions, and of the total errors, and average these; or we may take *twice* the percentile of the number of total errors and divide the sum by four, thus giving double weight to the errors made by the subject as against the time taken and the number of repetitions. Any other weighting of these three factors is, of course, possible. In considering errors, moreover, we may simply consider the unclassified errors, or we may give extra penalty for logical and perseverative errors, according to any one of several means conceivable. Each of these methods will give ratings of subjects slightly different from those derived by the others. But the differences are, after all, not as great as might be expected; for more time also means more errors and more repetitions, as a rule; on the whole, high ranking on any one of the criteria means high ranking on the others. The correlation between rankings by unclassified errors and total errors is but very slightly below 100. I got .99 with 81 subjects, the total error score being slightly better than the unclassified error score. The correlation is also very high between ratings in the Rational Learning Test when the total errors are counted on the one hand, and when total errors, minutes, and repetitions are considered on the other. In this case with twenty-nine subjects recently tested I got a correlation of .88, the better rating being obtained by the three-criteria method. While it is hardly possible at present to arrive at even a practical final judgment as to which criteria in the Rational Learning Test are most serviceable for the measurement of general intelligence, the best method appears to be simply to get the average of the percentiles of total errors, number of minutes, and number of repetitions. This measure is easily derived by the use of our percentile table. For example, the record of L. D. in Table I shows 144 total errors, 12 minutes, and 7 repetitions. The corresponding percentiles, as found in Table II, are: 65, 75, and 90, respectively. We simply take the approximate percentiles, the nearest multiple of five, as the tables are not yet accurately enough standardized to warrant taking the extra time for finding the exact percentile. Even to take the nearest percentile shown in the table is justifiable by the fact that the extremes—above 80 and below 20—are less accurate than are those nearer the median. R. L.'s percentile rank is therefore $\frac{65+75+90}{3}=77$. The general equation is

$$PR=\frac{P_{te}+P_{m}+P_{r}}{3}$$

PR is for percentile rank; P, percentile; and the small letters in the numerator, *te*, *m*, and *r*, are for total errors, minutes, and repetitions, respectively.

But in individual diagnosis it is not well to stop with one's *general* rating. We see that R. L. is ranked in the upper 25 per cent of college students, but in unclassified errors he surpassed only 65 per cent of college students, in logical error and perseverative error scores, 70 and 35, respectively. He is considerably below the median in perseverative errors. This probably indicates a subjectivity of mind, further data regarding which should be sought in the notes on his general behavior during the test, and in the detailed record of his responses, as shown in Table I. He collected himself very readily after confusions in the third series, and in the fourth made a perfect record, though two errors were again made in the fifth series. Notwithstanding confusions here and there, probably due to subjectivity of mind, R. L. seems to fall with the more logical, as opposed to the trial-and-error type of individuals.

The rank of a group can be obtained by finding the median percentile rank (PR) of all the members, and the various methods of comparing groups may be employed if the PR of each individual is taken as his score.

Other forms of the Rational Learning Test with ten letters may be obtained by making up a different order of the numbering of the letters, but in doing this care should be taken not to follow any subjective order that may be of aid to the subject. The norms here given apply only to the numbering shown in Table I, though they may be found to be fairly accurate for other forms. We have found no trouble with coaching thus far, but intend to standardize the test in at least two forms for different ages. This will not only aid in precautions against coaching but will add materially to the value of the test for various purposes. It appears desirable to use fewer letters in the test for children of the grades, as the present form is too hard for them. The median time for adults on the present test is slightly over sixteen minutes, and the determination of the subject's percentile rank after the test requires but a very short time.

If the test proves to be sufficiently worth while to warrant finding the subject's IQ, as used by Terman, this measure can readily be determined from the use of a table such as is given in the Otis' Group Intelligence Scale, *Manual of Instructions*. One must exercise caution in attempting to work out IQ's from standards derived from tests of college students, who compose a rather highly selected group.

The Rational Learning Test has been used in modified forms by Sunne, to whose report we have already referred, and by B. F. Haught. Sunne had twelve keys numbered in a random order, and so arranged that when key 1 was pushed a lamp would light; when this had been located, key 2 would next light the lamp; and so on. In this form spatial and other factors enter making results hardly comparable with those which we have presented here. Mr. Haught's method is to have ten keys under each of the ten letters we have used, one key under each letter being connected with a battery so as to ring a bell if pushed. No results by this method have yet been published. It is desirable in any such scheme as this to have some sort of self-recording apparatus; and it seems to be within the range of probability to have devised in time a machine with which the subject can be left wholly alone, and which will record in a convenient form, and in order of occurrence, all his significant responses. In the making of such a contrivance careful consideration should be given to the possible operation of spatial factors, as opportunity for space habits will reduce the difficulty of the problem for most persons. The present form of the experiment, however, whether used as a test or as a study of a higher form of learning, is not only convenient and highly objective in procedure so far as a study of the higher thought processes is concerned, but it also eliminates rather successfully the influence of spatial habits.

A SHORT POINT SCALE FOR MENTAL MEASUREMENT

By Esther Reedy and James W. Bridges, Ohio State University, Columbus

The increasing demand for intelligence ratings of large numbers of children and adults has resulted in the development of group examinations which are of undoubted validity and sufficient accuracy for the great majority of cases. However, these group examinations cannot yet entirely replace individual examinations. The latter are still necessary in the more careful diagnosis of subnormal, feeble-minded, and psychopathic cases. It seems possible however to reduce greatly the time required in giving such examinations. This could be done by selecting from one of the standard scales the tests of greatest diagnostic value and creating therefrom a short scale more or less comparable in reliability to the original. The present study was undertaken with the object of producing such an abbreviation of the Yerkes-Bridges Point Scale.

The data for the study were Point Scale examinations of five hundred and forty-one children between the ages of four and fourteen years. Two hundred and forty of these children were examined in the Psychological Clinic of The Ohio State University. The remaining three hundred and one were examined by L. E. Coler in two Columbus schools, one of superior and the other of inferior social status.[1] The chronological age and Point Scale score distributions of five hundred and forty-one cases are given in Table I.

The selection of tests for the short scale was based upon the following principles: correlation of test scores with whole scale scores, practical suitability, and relative difficulty. The relation of test scores to whole scale scores is indicated roughly in Table II which gives the median scores on the entire scale for subjects who make each score on each test. This table should be read in the following way: the median Point Scale score of subjects who make zero on test 1 is 40 points, the median Point Scale score of subjects who make 3 on test 5 is

---

[1] Bridges, J. W., and Coler, L. E.: The Relation of Intelligence to Social Status, *Psych. Rev.*, 24, 1917, pp. 1-31.

TABLE I

Chron. Age	No. Cases	Point Scale Score	No. Cases
4- — 4- 5	2	0- 4	0
4-6— 4-11	3	5- 9	1
5- — 5-5	3	10- 14	2
5-6— 5-11	9	15- 19	4
6- — 6-5	33	20- 24	15
6-6— 6-11	66	25- 29	26
7- — 7-5	38	30- 34	14
7-6— 7-11	46	35- 39	29
8- — 8-5	37	40- 44	45
8-6— 8-11	45	45- 49	57
9- — 9-5	30	50- 54	76
9-6— 9-11	24	55- 59	75
10- —10-5	24	60- 64	58
10-6—10-11	25	65- 69	45
11- —11-5	31	70- 74	39
11-6—11-11	27	75- 79	34
12- —12-5	27	80- 84	9
12-6—12-11	26	85- 89	3
13- —13-5	22	90- 94	7
13-6—13-11	23	95-100	2
Total	541		541

52 points, and so on. The amount of correlation was determined roughly by inspection of Table II. The range of the median scores in each test and their uniform increase as the test scores increase are the significant factors.

Certain tests which fulfill the correlation criterion were rejected on other grounds. Test 6 was omitted because it measures the same function as test 4, immediate memory; and test 4 has the advantage of uniform gradations in scoring. Similarly tests 10 and 19 measure ability to define. Test 10

TABLE II

Median Point Scale Scores of Subjects Making Each Score on Each Test

Test	*1*	*2**	*3*	*4**	*5*	*6*	*7*	*8*	*9**	*10**	*11*	*12**	*13*	*14*	*15**	*16**	*17**	*18*	*19*	*20**
Score																				
0	40	17	—	25	29	25	5	42	24	32	46	28	36	45	33	42	42	46	48	38
1	28	30	46	27	36	32	—	50	28	25	52	38	42	—	32	52	51	—	50	54
2	30	38	40	56	44	54	5	59	38	24	56	50	58	58	46	60	58	51	60	58
3	56	50	56	56	52	—	34	—	48	40	52	58	68	—	54	64	61	—	52	70
4	—	58	—	62	60	72	36	—	52	48	—	66	74	68	58	68	70	70	72	72
5	—	—	—	68	—	—	42	—	58	58	—	—	—	—	72	—	78	—	82	83
6	—	—	—	—	—	82	54	—	68	58	—	—	—	—	64	—	—	76	88	94
7	—	—	—	—	—		59	—	—	64	—	—	—	—	74	—	—	—	—	—
8	—	—	—	—	—		66	—	—	68	—	—	—	—	82	—	—	—	—	—
9	—	—	—	—	—		70	—	—	—	—	—	—	—	—	—	—	—	—	—

*Tests selected for Short Scale.

was given the preference because the selection of test 19 would tend to overweight the short scale with difficult tests, while it is obviously necessary to keep the short and long scales comparable in difficulty throughout the entire range. Test 7 was rejected because of its inability to discriminate in the lower levels of the scale, and also because of the personal equation in scoring. Test 13 was rejected because the time required for giving it (three minutes) does not correspond to its relative value in a short scale. The tests finally selected for the short scale are given below together with the maximum points credit for each test and the number of each test in the original Point Scale.

	Test	Points
2.	Missing parts..................	5
4.	Memory span for digits.........	5
9.	Comparison of objects..........	6
10.	Definition of concrete terms.....	8
12.	Copying square and diamond....	4
15.	Comprehension of questions.....	8
16.	Drawing designs from memory..	4
17.	Criticism of absurd statements....	5
20.	Analogies......................	5
	Total short scale score.........	50

The scores of the five hundred and forty-one subjects on this short scale were determined and compared with their scores on the entire scale. This comparison is shown in the correlation array, Table III. The Pearson product-moment coefficient of correlation is .95, which indicates that this short scale of nine tests is a fairly accurate measure of those functions measured by the entire scale. Correlations between chronological ages and scores were also computed. In the case of the entire Point Scale the coefficient was found to be .67, and in the case of the short scale, .64.

The relation of long scale scores to short scale scores was found to be approximately two to one. This is shown by inspection of the correlation array, Table III. It is also shown by comparison of percentile ranks. For example, the

TABLE III

THE CORRELATION BETWEEN SCORES ON THE ORIGINAL POINT SCALE AND SCORES ON THE SHORT POINT SCALE

Point Scale Scores	Short Scale Scores (0–1, 2–3, 4–5, etc.)																									
	0	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	Total
95–100																						1			1	2
90– 94																							2	4	1	7
85– 89																						2	1			3
80– 84																			1	1		4		3		9
75– 79																1		4	6	7	11	2	3			34
70– 74															1	1	3	7	9	12	5	1				39
65– 69														1	1	11	11	10	10	1						45
60– 64													2	2	9	21	7	11	4	2						58
55– 59												1	6	18	20	18	8	3	1							75
50– 54												15	20	20	13	7	1									76
45– 49										1	9	12	24	9	1	1										57
40– 44										4	9	18	11	1	2											45
35– 39									6	9	12	2														29
30– 34							1	5	6		1	1														14
25– 29						6	11	9																		26
20– 24				1	7	5	2																			15
15– 19		1	1	1	1																					4
10– 14	1		1																							2
5– 9			1																							1
0– 4																										
Total	1	1	3	2	8	11	14	14	12	14	31	49	63	51	47	60	30	35	31	23	16	10	6	7	2	541

seventy-five percentile rank is 66 points for the long scale, 34 for the short; the fifty percentile rank is 56 points for the long scale, 28 for the short; and the twenty-five percentile rank is 45 points for the long scale, 23 for the short. The score on the short scale can consequently be multiplied by two, and the result interpreted in terms of mental age by the standard Point Scale norms.

The short scale scores of the five hundred and forty-one subjects were multiplied by two and compared with their long scale scores in order to determine the deviation of the former from the latter. The median deviation disregarding signs is 3.9 points, but a deviation of 2 or less was found in two hundred and five cases. A distribution of the deviations over the whole range of the scale taking signs into account shows that the deviations tend to be minus in the ranges from 0 to 29, zero from 30 to 69, and plus from 70 to 100. Comparing these ranges with the standard age norms of the original long scale it seems that in determining mental ages from 6 to 12 years the short scale does not deviate greatly from the long scale, but that considerable deviations are found in the lower and higher ranges.

CONCLUSIONS

An abbreviation of the Point Scale consisting of nine tests with a total score of fifty points has been presented. This short scale can be used in schools, hospitals or elsewhere whenever time does not permit the use of the complete scale. The sum of the scores on these nine tests can be multiplied by two and interpreted by the usual Point Scale norms. If the result is not deemed satisfactory the subject may afterwards be given the remaining tests of the original scale. This will merely involve a slight change in the order in which the tests of the scale are given.

Comparison with the complete scale seems to show that this short scale gives fairly reliable results within the limits of six to twelve years. Scores on the short scale give a high correlation with scores on the complete scale, and the former correlate almost as high as the latter with chronological ages.

A STANDARDIZATION AND WEIGHTING OF TWO HUNDRED ANALOGIES

By RUDOLPH PINTNER, Ohio State University, and SAMUEL RENSHAW, Western State Normal School, Michigan

Of all the various forms of controlled association tests the mixed relations or analogies test is known to be one of the most useful and reliable. It differs from certain other association tests in that the form of the response is restricted. This restriction varies from one stimulus to another within the series of terms, and its nature is not revealed to the subject in the instructions but is supplied by the material and must be perceived by him from the material (6). Its widespread use in some form or other, has seemed to justify a somewhat more extensive weighting and standardization than has been attempted up to this time.

The form of the analogy seems to have had its origin with the Aristotelians, and it has come down to us through logic and formal grammar. Among the earliest attempts to use the analogous relation, its perception and completion, as a mental test was that of Woodworth and Wells (1). Similarly it was used by Whitley (2) in her study of the validity of various forms of association tests in respect to practice effect, variability and intercorrelation; by Burt (3) who gave the name analogies to the test; by Briggs (4) in his study of the disciplinary effects of instruction in grammar on the higher mental processes of school children; by Wyatt (5) in comparing abilities of Manchester children of various ages; by Bingham and Whipple (6) with American college students. More recently the test has been employed by Whipple, Henry, Manuel and Coy (7) for the selection of gifted children, by Rogers (8) in her tests for the diagnosis and prognosis of mathematical abilities, in the Otis Absolute point scale (9) and the Otis Group Intelligence test; in the Army Alpha tests; by Thurstone in his tests for college Freshmen and High School Seniors; by Yerkes, Bridges and Hardwick (14) in their Point Scale; by Thorndike (the pictorial and geometric form of the test) in his Non Verbal Intelligence Scale (15); by the Bureau of Salesmanship Research, and in many "omnibus" test groups.

Woodworth and Wells (1) claim the test measures "skill in handling associations," "flexibility of mental performance;" Burt (3) "perceptions, implicit or explicit, of the relation and reconstruction of the analogous one by relative suggestion;" Briggs (4) "rapidity and accuracy of mental reorientation," the "functioning of meaning tendencies." Whipple (7) is "confident that the analogies test brings out an ability that is decidedly symptomatic for the purpose of selecting gifted children." Wyatt (5) and Bickersteth (10) speak of it as a test of discrete reasoning.

Until the comparatively recent development of group testing, the lists of Woodworth and Wells, and of Whipple and Fraser, comprising some 20 single analogies to each list, were most commonly used, and this most frequently as an individual test wherein the association reaction time played a very important part.

There are two possible ways of scoring analogies:

1. To give credit according to the time required to find the analogy;
2. To give credit according to the difficulty of the analogy as determined by a standardization based upon the percentage of passes or failures.

Green (11) found such a high correlation ($r = .87$) between time as a measure of difficulty and the percentage of passes and failures, that it would seem best to follow the second of these methods. In general the scheme of weighting outlined by Rugg (12 and 13) was the one adopted in this study.

In making up our list of analogies we made selections from every available existing list, picking out the best ones and avoiding duplications as far as possible. A few original ones were added. After trying out in a preliminary way a large number, we finally selected two hundred and had them printed on a four-page folder, with fifty on each page. Ten point type was used for the first three terms, followed by four or five "choice" words, printed in bold-face type of the same size. A specimen of the original list is shown below:

51 better—good :: worse—**very-good medium bad much-worse best**
52 grass—cattle :: bread—**butter flour milk man horses**
53 large—object :: loud—**soft small heavy weight sound**
54 king—kingdom :: president—**vice-president senate republic queen democrat**

55 Man—revolver :: bee—**wings honey flying wax sting**
56 egg—bird :: plant—**seed shell leaf root feathers**
57 ignorance—education :: poverty—**laziness school wealth charity teacher**
58 circle—square :: sphere—**circumference cube round corners ball**
59 point—line :: line—**surface pencil dot curve solid**
60 disease—sanitation :: accident—**doctor hospital bandage cleanliness care**
61 ordinary—exceptional : : many—**all none few common more**

The subject is required only to underscore one of the choice words in black type to complete each analogy. Otis (9) has called attention to the fact that the chance-success factor in this procedure is more than atoned for by its ease of administration and scoring. Retrial with some of our subjects substantiates this view. No attempt was made in the limits of this study to deal with the pictorial or geometric forms of the test, devised and used by Thorndike (15) and by the National Research Council tests. Neither did we attempt, except in a superficial way, to try the construction of new and original analogies as a test. These are problems by themselves which would require special study and research.

The instructions to the subjects, after the usual preliminaries and samples, were in part: " You are to discover in what way some word among those in black type has the same relation to the third word in the series as the first and second have to each other. Then draw a line under that word. If you do not see the relation in about 8 seconds or are in doubt about it, mark the word that seems nearest it and go on to the next." Sufficient time was permitted for the subjects to mark each of the two hundred. The work limit method was employed with the first 100 subjects. The mean time for this group turned out to be about 26.5 minutes for the list. About 7 seconds is ample time to allow for each analogy in building a scale of the tests. It was found that if the subject was unable to perceive the relationship in about 4 or 5 seconds the chances were extremely great that never would he be able to do so correctly at that sitting. One group of 50 second-year College students were asked after completing the list to go back over it and carefully check the ones in their judgment regarded to be certainly right, wrong or doubtful. After scoring the papers it developed that the subject's opinion regarding the correctness or incorrectness of his responses has about a one-to-one chance of being right or wrong and so must be considered as uniformly unreliable.

The subjects numbered 917 in all. Of these there were 47 graduate students, 100 University Seniors and Juniors, 590 Normal College first and second-year students, and 180 High School students in grades XI and XII. The chronological ages ranged from 16 to 38 years. To determine whether we had a fairly random sampling of intelligence in our subjects we took a group of 52 second-year Normal College students and tested them with the Otis scale, form A. The scores ranged from 87 to 198, with the median at 158, Q_3 at 172, Q_1 at 151, $Q = 10.5$ and $\sigma = 6.7$. A product moment coefficient of correlation, uncorrected for attenuation, between the Otis and analogies scores for these subjects showed $r = .785 \pm .035$.

The ability measured by the analogies test shows rather steady and progressive increase with age. Our results shown in Table I conform fairly well to those of Bickersteth (10).

TABLE I

Group	Number	Mean Errors Of 200	Range
Graduate Students	47	21.8	5–59
Normal College Second Year	120	30.5	6–58
Normal College First Year	82	36.3	8–142
High School Seniors	97	35.0	11–153
*VIII. Grade	31	80.0	23–200

*Not included in our group of subjects.

Wyatt (5) found the test to correlate highly (.80) with intelligence. The reliability of the test as shown by him for 409 Oxford girls was:

Age	r
10	.35±.09
11	.62±.06
12	.78±.04
13	.76±.05

Bickersteth (10) adds "there does not seem to be much difference as the result of practice in the test once its nature is grasped by the subject—the marks gained in series 2 and 3 were often identical."

Burt (3) concludes that "those tests involving higher mental processes such as Reasoning, vary most closely with intelligence, and are least vitiated by various irrelevant conditions such as sex, social status, training of the experimenter, and mass measurements of numbers of subjects at once."

The weights of each analogy were determined by ascertaining the total number of the 917 subjects who failed each analogy. This number was then converted into a per cent of the total, by dividing by the number of subjects. Assuming a normal distribution of the trait measured in a curve whose limits are included in a range of $\pm 3\sigma$, setting 0 at—3σ, 50 at the mean and 100 at $+3\sigma$, the percentages of failures were transmuted into equivalent values of sigma. By the use of Rugg's table (12, appendix, Table VI) each was likewise assigned a percentile value.

In addition, each analogy was assigned a point value. This point value is purely arbitrary, and calculated in order to arrive at groups of analogies which are roughly of equal difficulty. The sum of all the percentile values of the 200 analogies is 6,002. Dividing this by 1,000, we get a constant 6 by which all the percentiles have been divided, thus turning the percentile values into relative point values. The sum of all the point values of the 200 analogies is 1,000.

In this way we have three possible methods of weighting or scoring each analogy. The sigma values give the finest differentiations and lend themselves to as fine a system of scoring as may be desired. The percentiles give a somewhat rougher method, and the point values still more so. Using either of these three methods we may easily construct different lists of analogies of equal difficulty.

We print here the list of 200 analogies arranged according to difficulty, giving the percentile, sigma and point values. The correct response is italicized.

POINT VALUE

Percentile	Sigma	
		10.0
62	3.76	heat—a gas :: cold—ice winter *water* refrigerator
61	3.68	truth—falsehood :: pride—fear shame honor *humility*
60	3.64	toil—soil :: pay—check *gay* debt money
	3.64	darknes—sunlight :: stillness—quiet *sound* loud dark moonlight
		9.5
58	3.53	point—line :: line—*surface* pencil dot curve solid
	3.49	oil—toil :: hate—love work boil *ate* hat
57	3.47	motion—physics :: blood—temperature body veins *physiology* geography

9.0

55....3.34 sorrow—misfortune :: joy—grief happiness hatred *success* pride

3.32 anger—violence :: love—*caressing* hate temper hope happiness

3.31 circle—ellipse :: square—cube curve oval circle *diamond*

3.31 land—wealth :: sun—*life* earth death moon

3.31 evolution—revolution :: crawl—baby floor stand *run* hands and knees

8.5

52....3.17 fear—anticipation :: regret—*memory* hope sorrow hate forget

51....3.11 book—knowledge :: money—paper dollars *bank* work gold

50....3.00 Sunday—week :: January—*year* hour Wednesday month

8.0

49....2.99 moon—earth :: earth—Mars *Sun* clouds stars universe

2.97 land—peninsula :: ocean—river lake cape *gulf* water

2.97 known—unknown :: present—past expensive *future* Christmas

7.5

47....2.87 tolerate—pain :: welcome—*pleasure* unwelcome friends give

2.86 better—good :: worse—very-good medium *bad* much-worse best

2.84 disease—sanitation :: accident—doctor hospital bandage cleanliness *care*

46....2.77 electricity—wire :: gas—flame spark hot *pipe* stove

2.76 engineer—chauffeur :: locomotive—iron stack engine *auto*

2.76 character—important :: complexion—blonde *trivial* brunette good

2.76 behind—late :: before—after soon *early* dinner

45....2.70 advice—command :: persuasion—help aid urging *compulsion*

7.0

43....2.62 man—skull :: church—umbrella rain Sunday *dome*

2.61 ice—water :: water—land *steam* cold river thirst

2.61 arteries—body :: railroads—*country* train crossing accident

2.60 beautiful—appearance :: sweet—*taste* beauty sour ugly nice

2.60 uncle—aunt :: son—brother *daughter* sister father girl

42....2.57 city—mayor :: army—soldier navy private *general*

2.57 officer—private :: command—army general *obey* regiment

2.52 camp—safe :: battle—win field fight *dangerous*

6.5

41....2.48 wood—table :: knife—cutting chair fork *steel* handle

2.46 rudder—ship :: tail—*bird* sail dog cat

40....2.44 quarrel—enemy :: agree—policeman agreeable foe *friend*

2.40 ordinary—exceptional :: many—all none *few* common more

2.40 large—elephant :: loud—soft hear *cannon* see

2.40 lion—animal :: rose—small bird *plant* thorn

39....2.38 city—telephone-system :: body—arteries *nerves* arms clothes skeleton

6.0

38....2.33 ignorance—education :: poverty—laziness school *wealth* charity teacher
2.30 circle—square :: sphere—circumference *cube* round corners ball
2.29 imitate—copy :: invent—study invention machine *originate*
37....2.27 sand—glass :: clay—stone hay *bricks* dirt
2.25 liquid—solid :: water—salt steam lake *ice*
2.25 succeed-fail :: praise—lose friend God *blame*
36....2.19 large—object :: loud—soft small heavy weight *sound*
2.17 wolf—sheep :: cat—fur kitten dog *mouse*
2.16 rafters—house :: skeleton—bones skull grace *body*

5.5

35....2.15 cold—ice :: heat—wet cold *steam* stars
2.15 success—joy :: failure—*sadness* success fail work
2.15 past—present :: yesterday—*today* tomorrow Christmas gone
2.15 writer—books :: bee—hive *honey* wasp sting
2.15 mountain—valley :: genius—*idiot* right think brain
2.13 tiger—carniverous :: horse—cow pony *herbiverous* buggy
2.13 tears—laughter :: sorrow—*joy* distress funeral sad
2.12 part—whole :: Indiana—state Ohio Indianapolis *United States*
2.11 hope—despair :: happiness—frolic fun joy *sadness*
2.11 cold—heat :: ice—cream frost refrigerator *steam*
34....2.09 Sunday—Friday :: Monday—Tuesday *Saturday* week Thursday
2.09 eraser—ink :: water—lightning storm *dirt* clothes
2.07 dismal—dark :: cheerful—fun *bright* house gloomy
2.06 grass—cattle :: bread—butter flour milk *man* horses
2.06 egg—bird :: plant—*seed* shell leaf root feathers
2.05 breeze—cyclone :: shower—bath flood winter spring
2.04 lion—animal :: rose—smell leaf *plant* thorn
33....2.02 Japanese—Japan :: Dutch—Russia *Holland* Siberia Spanish
2.02 prosperity—happiness::adversity—success *sorrow* fun rage
2.00 arm—elbow :: leg—foot *knee* stocking toe heel
2.00 hunter—gun :: fisherman—fish bold wet *net*
1.98 music—noise :: harmony—hear accord violin *discord*

5.0

32....1.96 king—kingdom :: president—vice-president *republic* queen democrat
1.96 bird—song :: man—woman *speech* boy work
1.96 peace—happiness :: war—*sorrow* fright battle Europe
1.93 horse—mule :: obedient—disgraceful *stubborn* donkey obey
31....1.88 electric light—candle :: automobile—*carriage* electricity tire speed glow
1.88 truth—gentlemen :: lie—*rascal* live give falsehood
1.87 ruler—length :: clock—hour distance *time* alarm
1.86 head—hat :: hand—arm *glove* finger wrist
1.86 man—boy :: sheep—wool *lamb* goat shepherd dog
1.86 peninsula—continent :: bay—Massachusetts boats pay *ocean*

30....1.85 pretty—ugly :: attract—fine *repel* nice draw
1.84 pupil—teacher :: child—*parent* doll youngster obey
1.84 hospital—patient :: prison—cell *criminal* bar jail
1.82 wash—face :: sweep—broom *floor* straw clean
1.82 book—writer :: statue—liberty *sculptor* picture state
1.82 captain—ship :: mayor—state council *city* ship boss

4.5

29....1.77 eat—fat :: starve—food *thin* bread thirsty
1.76 Monday—Tuesday :: Friday—week Thursday day *Saturday*
1.76 man—revolver :: bee—wings honey flying wax *sting*
1.76 pan—tin :: table—chair *wood* legs dishes
1.75 body—food :: engine—wheels *fuel* motion smoke fire
1.75 tears—sorrow :: laughter—girls grin *joy* sob
1.74 lead—bullet :: gold—paper *coin* silver copper
28....1.71 picture—see :: sound—noise music bark *hear*
1.71 lobby—hotel :: preface—author *book* porter elevator
1.70 water—fish :: air—spark *man* blame breathe
1.70 man—arm :: tree—shrub *limb* flower bark
1.69 spoon—soup :: fork—knife plate *meat* cup
1.68 floor—ceiling :: ground—earth hill grass *sky*
27....1.67 Edison—phonograph :: Columbus—Spain *America* Washington Ohio
1.66 poison—death :: food—eat bird *life* bad
1.66 singer—song :: preacher—choir organist pulpit *sermon*
1.65 abundant—cheap :: scarce—buy bargain *costly* nasty
1.64 floorwalker—store :: policeman—fire *street* conductor wagon
1.64 peeling—banana :: shell—skin orange *egg* juice ripe
1.64 foot—man :: hoof—leather shoe *cow* leg
1.63 cannon—large :: rifle—ball *small* bore shot
1.63 historian—facts :: novelist—*fiction* Dickens writer book
1.63 bold—timid :: advance—proceed *retreat* campaign soldier
1.63 ear—hear :: eye—hair blue *see* eyebrow.

4.0

26....1.60 razor—sharp :: hoe—bury *dull* cuts free
1.60 light—dark :: noise—report ring *silence* sound
1.59 book—author :: statue—*sculptor* marble model magazine man
1.59 establish—abolish :: begin—work year *end* commence
1.59 ocean—pond :: deep—sea well *shallow* steep
1.58 flag—country :: cross—purgatory *christianity* army president
1.58 December—January :: last—*first* least worst month
1.57 boy—man :: lamb—wool dog *sheep* shepherd
1.57 order—confusion :: peace—part treaty *war* enemy
1.57 hat—head :: thimble—*finger* needle thread hand sewing
25....1.55 reward—hero :: punish—God everlasting pain *traitor*
1.55 hand—arm :: foot—*leg* toe finger wrist elbow
1.53 Yes sir—No sir :: always—meanwhile however perhaps *never*
1.53 confusion—order :: war—guns *peace* powder thunder army
1.52 engine—caboose :: beginning—commence cabin *end* train

1.52 10—100 :: 1000—money *10,000* 20,000 wealth
1.52 knitting—girls :: carpentry—trade houses *boys* lumber
1.52 add—subtract :: multiply—add *divide* arithmetic increase
1.52 *automobile—wagon :: motorcycle—ride speed bicycle car*
1.51 foot—man :: hoof—leg dog *horse* boy shoe
1.51 complex—simple :: hard—brittle money *easy* work
1.51 frame—picture :: lake—porch photo window *island*
24....1.48 sweet—sugar :: sour—sweet bread man *vinegar*
1.47 a—b :: c—e *d* b letter
1.46 left—right :: west—south direction *east* north
1.46 granary—wheat :: library—desk *books* paper librarian
1.46 dog—bark :: cat—fur chase *mew* mouse
1.44 palace—king :: hut—*peasant* barn farm city

3.5

23....1.43 man—woman :: brother—daughter *sister* boy mother son
1.43 door—house :: gate—chimney *yard* swing window
1.41 hope—happiness :: despair—grave repair death *grief*
1.41 sit—chair :: sleep—*bed* rest wake snore
1.38 winter—summer :: cold—freeze *warm* wet January
22....1.37 blonde—brunette :: light—heavy electricity *dark* girl
1.35 *esteem—friends :: despise—enemies forsake detest people*
1.35 Yes—No :: affirmative—win debate deny *negative*
1.35 cellar—attic :: bottom—well tub *top* house
1.35 birth—death :: planting—corn spring *harvest* wheat
1.34 telephone—hear :: spyglass—shout telegraph distance *see*
1.34 angels—heaven :: men—*earth* women boys paradise
1.34 engineer—engine :: driver—*horse* harness passenger man
1.32 polite—impolite :: pleasant—*disagreeable* agreeable man face
1.32 prince—princess :: king—palace *queen* president kingdom
1.32 wool—sheep :: feathers—pillow rabbit bird goat bed
21....1.30 egg—bird :: seed—crack *plant* grow nest
1.30 skin—body :: bark—*tree* dog bite leaf
1.26 clock—time :: thermometer—watch warm bulb mercury *temperature*

3.0

1.25 skating—winter :: swimming—diving floating hole *summer*
1.22 hour—minute :: minute—man week *second* short
1.22 Washington—Adams :: first—contrast best second last
20....1.20 cannon—rifle :: big—*small* bullet gun army
1.20 parents—command :: children—order shall *obey* must
1.16 giant—dwarf :: large—dog monster queen *small*
1.16 1—2 :: 3—4 2 *6* 9
19....1.14 man—home :: bird—*nest* fly insect free
1.14 corn—horse :: bread—daily flour *man* butter
1.14 wool—sheep :: fur—*cat* bird hat coat
1.11 good—bad :: sweet—taste conduct *sour* polite
1.11 roof—house :: hat—shoe straw attic *head*
1.11 go—come :: sell—leave buy money paper
18....1.09 roof—house :: hat—button shoe straw *head*
1.09 wrist—cuff :: neck—leg giraffe foot *collar*

2.5

17....1.07 suitcase—clothing :: purse—purchase *money* string stolen
1.07 *Washington—Adams :: first—last worst *second* best
1.07 able—unable :: strong—muscle exercise axe *weak*
1.00 sailor—navy :: soldier—gun cap hill *army*
1.00 pitcher—milk :: vase—*flowers* pitcher table pottery
1.00 white—black :: good—mother time clothes *bad*
16....0.98 uncle—nephew :: aunt—*niece* brother sister cousin
0.98 December—Christmas :: November—month *Thanksgiving* December early
0.98 terrier—dog :: Jersey—city *cow* horse state
0.98 straw—hat :: leather—feather cool soft *shoe*
0.94 airplane—air :: submarine—dive engine ship *water*
0.94 dog—puppy :: cat—tiger dog horse *kitten*
15....0.90 *locomotive—train :: horse—station hut baggage *buggy*
0.90 dress—woman :: feathers—neck feet bill *bird*
0.90 food—man :: gasoline—gas oil *automobile* spark

2.0

14....0.86 above—below :: top—spin *bottom* surface slide
0.86 table—wood :: stove—bottle paper *iron* cork
0.86 skirts—girl :: trousers—hat vest *boy* coat
13....0.82 abide—depart :: stay—over home play *leave*
0.77 locomotive—coal :: automobile—motorcycle smoke wheels *gasoline* burn
0.77 theatre—people :: hive—thrive sting *bees* thick
12....0.73 shoe—foot :: hat—coat nose *head* collar

1.5

11....0.66 handle—hammer :: knob—key *door* shut room
9....0.59 doctor—patient :: lawyer—nurse hospital court *client*
0.59 gun—shoot :: knife—run *cut* hat bird

1.0

6....0.40 cradle—baby :: stable—*horse* man dog cat

* Used as samples in the instructions to subjects.

BIBLIOGRAPHY

1. WOODWORTH, R. S. and WELLS, F. L.: Association Tests, *Psych. Rev. Mon.*, Vol. XIII, No. 3, Whole number 57, Dec., 1911.
2. WHITLEY, M. T.: An Empirical Study of Certain Tests for Individual Differences, *Archives of Psych.*, No. 19, 1911.
3. BURT, C.: Experimental Tests of the Higher Mental Processes and Their Relations to General Intelligence, *Jour. of Exp. Ped.*, Vol. I, 1911, pp. 93-112.
4. BRIGGS, T. H.: Formal English Grammar as a Discipline. *Teachers College Record*, Vol. XIV, 1913, pp. 251-343.
5. WYATT, S.: The Quantitative Investigation of Higher Mental Processes, *Brit. Jour. of Psych.*, Vol. VI, 1913, pp. 109-133.
6. WHIPPLE, G. M.: *Manual of Mental and Physical Tests*, Vol. II, Baltimore, 1915, pp. 89-94.
7. WHIPPLE, G. M. et al.: *Classes for Gifted Children*, 1919, p. 25 and p. 36.

8. Rogers, A. L.: Experimental Tests of Mathematical Ability and Their Prognostic Value, *Teachers College Contributions to Education*, No. 89, 1918.
9. Otis, A. S.: An Absolute Point Scale for the Group Measurement of Intelligence, *Jour. of Ed. Psych.*, Vol. IX, Nos. 5 and 6, May-June, 1918.
10. Bickersteth, M. E.: Mental Tests of Childen of Various Ages, *Brit. Jour. of Psych.*, Vol. IX, Dec., 1917, pp. 33 ff.
11. Green, H. A.: A Standardization of Certain Opposites Tests, *Jour. of Ed. Psych.*, Vol. IX, No. 10, Dec., 1918, pp. 559-566.
12. Rugg, H. O.: *Statistical Methods Applied to Education*, Houghton Mifflin, 1917.
13. Rugg, H. O.: Standardized Tests and the Improvement of Teaching in First Year Algebra, *School Review*, Feb. and Mar., 1917.
14. Yerkes, R. M., Bridges, J. W., and Hardwick, R. S.: *A Point Scale for the Measurement of Mental Ability.* Baltimore, 1915.
15. Thorndike, E. L.: A Standardized Group Examination of Intelligence Independent of Language. *J. of Appl. Psych.*, Vol. III, No. 1, March, 1919, pp. 13-32.

BOOK REVIEWS

L. M. Gilbreth. *The Psychology of Management.* Sturgis and Walton Company, New York, 1918, p. 344.

Any writer who claims that his book "will aid the cause of Industrial Peace," deserves a hearing at the present, when the whole world is afflicted with strikes and internal conflicts. Such a claim is made by L. M. Gilbreth, the author of *The Psychology of Management.*

This book was first published in 1914, reprinted two years ago without any revisions. Nevertheless the book still retains some peculiar features which characterize no other psychological work. They do not perhaps constitute any contribution to academic psychology, but they are entirely in harmony with the psychology of common sense and of everyday life. We refer to the various terms the author uses and the way in which he defines them. In no case has he attempted a new definition; but he has left no important terms undefined. Whenever possible, a dictionary meaning is given.

The presentation of the subject matter is deductive and very clear. This procedure has at least one distinct value: it considers popular conceptions as they are rather than as what they should be; hence the conclusions derived from his statements are applicable to practical business management, in so far as they are correct.

The whole book is divided into ten chapters. Each chapter, with the exception of the first one, considers an important topic, such as Standardization or Functionalization. Considerable space is given to the exposition and interpretation of Taylor's system of scientific management. Its development is traced through various stages. It is only in Chapter 2, under the heading of "Individuality," Chapter 9, under "Incentives," and Chapter 10, under "Welfare," that we find some psychological discussions. Even here the topics could be brought much more up to date. The term "instinct," which is so much in vogue now-a-days in the literature on social psychology, finds no consideration in this book, although it is mentioned here and there. One wonders how workers can be educated to the point "where they will be fitted to work and to live" and how "the cause of industrial peace" can be attained without at least a preliminary study of the fundamental human instincts.

The author, however, is not to be held responsible for this serious omission. His work is typical of its time and serves to illustrate the infantile stage of psychology as a science at the time of his writing. While he "aims at no exhaustive study of psychology" in relation to industry, his book may be considered a valuable contribution to this particular field.
C. LI.

W. B. Pillsbury. *Psychology of Nationality and Internationalism.* D. Appleton and Co., N. Y., 1919, p. 314.

The author here presents a keen analysis of the criteria of nationality, and their practical bearing on the problems of free nations. While the Versailles treaty provides that henceforth each nation may be self-determining, there appears to be no agreement as to what constitutes a nation. Before any national criteria can be evaluated, a distinction must be made between the concepts *Nation* and *State.* These expressions are commonly employed indifferently in reference to one or the other. "Nationality is the mental state or community in behavior," while the state "is merely the system of government, a unity for the sake of making and enforcing laws."

Upon such a premise, our author finds that the language criterion of nationality has no claim for consideration aside from what it contributes to individual expediency in becoming naturalized. Compared with the instinctive and emotional basis, which appear to offer the fundamental sources for nationality, language, race, history, and place of birth are relatively unimportant considerations. At most these must be regarded as secondary criteria, for they refer to nationality on a small scale. Nationality as it exists today, transcends all the characteristics which may properly enough belong to family and tribal organizations, even though such groupings appear before the nation as such. The popular genetic conception of nationality in respect to the family or tribe is no longer satisfactory.

A nation is a psychological unit, based upon certain primary instinctive tendencies found among all peoples. Upon these instincts acceptable ideals are evolved, pretty much as social and economic situations approve. Relativity is the big word in respect to both instincts and ideals. The chief instincts here involved are gregariousness, sympathy, respect for the opinions of others, taking on the nature of fear and hate. It would appear that fear alone is the chief generator of the countless ideals which complete the requisites for communal and national life in the psychological sense. While national ideals represent a gradual accumulation in most cases, they may be rapidly fostered through education. Herein lies our optimism for the future. Liberty and freedom have been the prevailing national ideals, except in the case of Germany; here the ideal was a super-personal state—the surrender of the individual to a mythical and unerring entity from which there could be no appeal. Whatever the national ideal may be, it, and the propensity of each individual to follow the same, constitute the two big factors in the consciousness of nationality.

An important question before the world today is whether or not nationality is the last word in political organization. Our author finds every psychological reason to believe that the final step in the federation of the nations of the world can and must be made. Were the problems of nationality confined to instinct alone, each nation as it exists would represent a closed unit. Happily the ideals which may be created from time to time point to the solution. These, when well established, have the force of instinct which insures unity and progress. "One may confidently assert that the development of the national spirit has come about by restriction of the natural range of the social instincts by training rather than by any unnatural extension of them."

Aside from the main thesis above stated, the book is replete with valuable suggestions of psychological and practical significance. For many people it must seem that race hatred is the ultimate cause of war, since it is regarded as a more primitive instinct than sympathy. Such a notion is denied by the author. Hate, while in itself is one of the strongest social forces operative at the present time, really developed in the interest of group survival, having sympathy as the ultimate impulse. There are therefore certain ambivalent tendencies in the individual and social minds which must not be overlooked. Again it is popularly supposed that morality in general seems to be weak among the children of foreign born parents, since they deplore the language, customs and admonitions of their parents. Yet when the situation is analyzed such tendencies make for American citizenship, even though something is temporarily lost to domestic felicity. From the viewpoint of emotional life we may judge therefore how far and in what direction democracy in America is doing its work. If we can keep our customs intact and appreciate adherence thereto, the alien alone will meet the fundamental requirements of citizenship.

The book as a whole exemplifies the tendency among modern psychologists to take a synthetic view of psychology. Both the emotions and intellect are here treated in their adequate relations. While it is not a war book in the popular meaning, it is none the less one of the best of them, for the author has considered the nations at war as his laboratory. The book requires careful study and should be read by everyone interested in the psychology of war and the issues that have grown out of it.

University of Vermont. J. W. SPROWLS.

NORAH H. MARCH. *Towards Racial Health.* E. P. Dutton and Co., New York, 1919, p. 320.

Miss March's book is evidently designed to serve a double purpose,—to give an insight into the sexual life of the child and to present a practical and efficient plan of sexual education for the use of parents and teachers. Incidentally, its chapters include many notes on sex hygiene which should be invaluable for the guidance of the mother in the care of her children.

The volume is divided into the following chapters: The Physical Development of the Child; The Mental and Emotional Development of the Child; Care of Children; Supervision—Psychological Aspect; Nature Study in the Service of Sex Instruction; Further Aids Toward Understanding the Biology of Sex; Ethical Training; Education for Parenthood; Education for Parenthood—Some Suggestions; Social Safeguarding.

There are very many commendable features in Miss March's treatment of her subject. She shows a thorough knowledge of the physiology of sex, and is able to put her information into surprisingly nontechnical terminology, although including such complicated processes as hormone action in her discussion. Her psychological deductions follow very closely those of G. Stanley Hall, as she indicates in her preface that they may do. She has dwelt very lightly on the pathological side of the sex life, which is probably quite as well, since that has been sufficiently emphasized of late by a number of specialists, so that we have almost come to forget the normal aspects of the problem.

To the student of psychology, Miss March's book offers no new information. But she has deliberately chosen to write an extensive rather than intensive study. Since her book is intended for an eminently practical purpose, this choice is justified. It is quite as necessary that the results of our scientific research be put into simple and comprehensive statements from time to time as that our investigations be continued. Only by such non-technical summaries can the results of research be made available for general use.

There is no doubt that Miss March has succeeded very ably in her task of popularization of sex physiology and psychology, and her book is to be recommended to parents and teachers both because of its clarity of expression and its frank and healthy attitude toward the subject with which it deals. PHYLLIS BLANCHARD.

ANDRÉ TRIDON. *Psychoanalysis, Its History, Theory and Practice.* B. W. Huebsch, New York, 1919, p. 272.

Tridon's book is really a masterly and comprehensive survey of the work thus far accomplished in psychoanalysis. It is written with clear, synthetic grasp, and keen insight into the problems presented by conflicting theories. The scope of the work is indicated by the chapter headings: The History of Psychoanalytic Research; The Unconscious and the Urges; The Repression of the Urges; Night Dreams and Day Dreams; Symbols, the Language of the Dream; The Dreams of the Human Race; The Psychology of Everyday Actions; Feminism and Radicalism; The Psychology of Wit; The Artistic Temperament; The Urges and Literature; The Urges and the Arts; Forms of Abnormal Compensation; The Oedipus Complex; The Neuroses. Epilepsies and Psychoses; Perversions; Crime and Punishment; The Psychoanalytic Treatment; The Transference; Re-education and Prophylaxis; The New Ethics.

Although Tridon lays no claim to attempting anything more than a summary of the contributions of the various schools to psychoanalysis, as a matter of fact he does far more than that. With impartial sureness, he analyzes and evaluates and places in proper perspective the work of Freud and his followers, of Adler, of Jung and the Zurich school, of Kempf, et als. Not that he consciously sets out to do this, but his treatment of the various subjects mentioned in his chapter headings incidentally involves such an evaluation. For example, he sees the "ego urge" of Adler and the "sexual urge" of Freud not as separate, all-important "urges" back of human behavior, but rather in their reciprocal relationship.

"We must bear in mind," he says, "that the three main urges, like all human phenomena, are closely related and can never be considered as absolute entities. Sucking the mother's nipple, which in the infant is primarily an activity to secure food, develops into a semi-sexual activity totally unrelated to nutrition and from which kissing originates. The physical pleasure a Don Juan derives from a new conquest increases his egotism and his sense of power; and reciprocally, Don Juan's newly acquired sense of power and increased egotism, revealed by certain attitudes, postures, buoyancy, mental and physical, may increase his sexual pleasures by vouchsafing him new conquests, etc." (p. 28.)

Of course, some of Tridon's chapters show a broader orientation and a more correct evaluation of the material than others. The section on repression of the urges, for instance, would be much more adequate were the work of Trotter (*The Instincts of the Herd in Peace and War*) and of Kempf (*The Tonus of the Autonomic Segments as Causes of Abnormal Behavior, J. Nerv. & Ment. Disease,* Jan., 1920) utilized in the explanation. It is quite true that the urges are often denied expression by the realities of the environment. It is equally true, as Trotter has pointed out, that a great many repressions are due to the desire to follow the leadership of the herd, and are only explicable as the inhibition of other "urges" (to use the author's term), by the gregarious impulse. Kempf expresses this very plainly when he writes:

"A most important factor begins to exert pressure upon the infant at birth and continues throughout its life. It is the incessant, continuous pressure of the herd . . . to conventionalize its methods of acquiring the gratification of its needs."

It may be that this view of the mechanism of repression is implied in Tridon's discussion, but it might well have been stated more clearly and unmistakably.

Most of the subjects are very ably handled, however. It may be that the author lays somewhat more stress upon the Adlerian theory than upon that of Freud, but it is an open question whether the facts at hand do not justify such an emphasis. The occasional original interpolations are so penetrating and show so much discernment that the reader cannot but wish that Tridon had indulged in more of these side remarks and interpretations of his own. In view of the writer's clearness of thought, and conciseness of expression, it is hardly sufficient to classify the book merely as the one broadly synthetic work in the field of psychoanalysis. Even on that basis, however, it should prove an important addition to psychoanalytic literature.

PHYLLIS BLANCHARD.

SUMMER H. SLICHTER. *The Turnover of Factory Labor.* D. Appleton & Co., New York, 1919, p. 460.

Throughout his thesis the author sets forth the urgent need for a definite and well-defined labor policy in industry. He speaks from his actual experience as a worker among workmen as well as from an analysis of facts gathered from many channels of recent research.

Employers do not have time for an analysis of the human element in the productive process from the point of view of the human being himself, so Slichter's problem is getting the employer to feel the things the worker experiences every day at his work. Many surprising reactions are shown which may be expected from workers under various conditions, under "panicky" and prosperous times.

How a feeling of estrangement between men and management, the "drive" system of management, the spirit of profit making, the social stigma attached to certain types of work, all tend to kill interest and take the soul out of the worthwhile activity, are only a few of the many ideas clearly treated in explaining the present attitude of the working class.

Following a keen analysis of the motivation of the workers' behavior, a constructive labor policy is outlined. The keystone of his policy is the Supervisor of Labor. The responsibility for handling rates of wages, working hours, promotions, amelioration of working conditions, grievances, and foreman training is concentrated in the hands of this executive to assure adequate attention. The type of man for this job is of supreme importance. "Only when the determination of the labor policy is placed in the hands of a Specialist in labor problems will the problem . . . be solved."

The extensive field which the material and references cover makes this a work on labor problems. Superintendents, employment managers, foremen, social workers and students of social psychology will find the book full of vital information in regard to the motivation of the individuals who daily pass in and out of the employment gate of a large factory.

Carnegie Institute of Technology, Pittsburgh, Pa. MYRON F. BARRETT.

HERBERT WOODROW. *Brightness and Dullness in Children.* J. B. Lippincott Company, Philadelphia, 1919, pp. 322.

The author calls this book an introduction to the science of general intelligence, and rightly so, for his language is so simple and the presentation of facts and theories so lucid and stimulating that it should prove intelligible as well as interesting to the reader who has no previous knowledge of a general psychology.

Without entering into a theoretical discussion of the various hypotheses concerning the nature of general intelligence, the second chapter takes up, in a somewhat chronological order, the various scientific attempts at the measurement of intelligence, admitting as the only valid objection to the Binet tests that "the influence of the tester's personality cannot be entirely eliminated" (p. 39). In the third chapter an important distinction is drawn between intelligence as measured by mental age or as an absolute amount which increases with age, and intelligence as "more or less constant throughout life" or as "brightness," of which idiocy and the superiority of genius represent the possible extremes. The term is used by the author in the first sense, with the further qualification that "intelligence means not the capacity for success along any one line of endeavor, but general capacity" (p. 148) for success in all those tasks and performances which require mental activity for their execution. Nor is any one special phase of such activity to be identified with intelligence, as various other writers have attempted to do. Our author is similarly guarded in his discussion of the intelligence quotient, which, he says, "cannot be expected to remain constant except for the average" (p. 52).

The topic of brain development in children is illustrated by an hitherto unpublished curve of Dr. Richard Scammon which he derived from a very careful study of all existing data. A short chapter on physical defects and their bearing on general intelligence is followed by a treatment of the problems of anatomical or physiological age in its relation to mental age which is superior to that of any other writer known to the reviewer. In the chapter on pedagogical age emphasis is laid upon the causes of and remedies for retardation in school. The next three chapters deal with the relation of intelligence to the more fundamental, simple and complex mental processes and capacities, leading up to the important problem of mental organization, the treatment of which is, unfortunately, too brief to equal other chapters in lucidity; yet it is along the lines here indicated that the author's greatest opportunity seems to lie for an important and lasting contribution to the science of general intelligence and to psychology. The chapter on heredity, with its illustrative case studies, prepares the way for the discussion of the problems and methods of special instruction of dull, feeble-minded, and superior children.

The most valuable feature of the book as a whole is the successful combination of scientific accuracy with simplicity of style and concreteness of subject-matter. The latter has been enhanced by well-selected illustrations, among which are two full-page radiographs showing carpal development at varying anatomical ages. L. R. G.

RANSOM A. MACKIE. *Education During Adolescence.* E. P. Dutton and Company, New York, 1920, pp. 222.

This book is intended as a general introduction to the principles of secondary education and is admittedly based very largely upon G. Stanley Hall's works on "Adolescence" and "Educational Problems."

President Hall has contributed a short Introduction to the book of his former student. The first chapter, entitled "Education During Adolescence," is a resume of Hall's general ideas on the subject, to which the author adds by way of "comment" four specific aims that high schools should pursue, relating to physical well-being, vocational guidance, personal culture, and social efficiency. The second chapter deals with the six-year high school curriculum, the third with principles of election. Here many opinions are quoted, almost at random, ending with the author's own opinion. The heading of the fourth chapter is more promising: "Changes Proposed in Secondary Education," but after a few more citations and some vague references to the Boston and Seattle schools we arrive at such generalities as this: "In order to be self-supporting, each individual should take studies bearing directly or indirectly on some trade, occupation, or profession. * * * But he should be trained for genuine citizenship. * * * and * * * he should secure 'general culture'" (p. 77). The remaining three chapters deal with the "Required Subjects," namely, Social Studies, English, and History, and consist almost entirely of loosely connected quotations from Hall, Elliott, and reports of committees of the National Education Association. With the present shortage of labor and material, this kind of literary hero-worship is greatly to be deplored, even if it were accompanied by valuable original offerings, as is not the case in the book before us. L. R. G.

W. J. Crawford. *Experiments in Psychical Science.* E. P. Dutton and Company, New York, 1919, p. 201.

The author of this volume is a Lecturer in Mechanical Engineering at the Municipal Technical Institute and at Queen's University, Belfast. The present book is in part a continuation of his previous work on *The Reality of Psychic Phenomena.* The author describes in some detail his methods and the results obtained with levitation, contact phenomena, and direct voice phenomena. He gives minute descriptions and measurements of all conditions, utensils, and other items involved; for instance, the weight of the medium before and after the seance, the weight of the levitating table as well as its dimensions, and even its temperature during a long levitation. He illustrates his explanations with many figures and drawings and four full-page photographs. On the basis of answers received from the "spirit operators" to questions about their methods of producing their psychic phenomena the author attempts to construct several hypotheses purporting to explain these phenomena analogous to the laws of mechanics. He believes, for instance, that "a psychic structure" emanates from the body of the medium built up of organic matter in a very unstable state and acutely sensitive to practically all light, and exclaims: "Imagine then the devastating effect of the magnesium flash upon this delicate structure." Nevertheless this same psychic structure in other respects obeys the laws of mechanics. The experiments with voice phenomena were practically without positive results. In spite of all pseudo-scientific efforts to introduce a few physical measurements into the seances the author does not seem to have been able to avoid self-deception, and his experiments are not convincing. L. R. G.

WALTER SCOTT MONROE. *Measuring the Results of Teaching.* Houghton Mifflin Company, Boston, 1918, pp. 297.

This book, which is intended for the teacher in the elementary school, describes a number of standardized educational tests in each of the following subjects: reading, arithmetic, spelling, handwriting, language and grammar, geography, and history. The emphasis is not laid upon detailed directions for giving the tests, but on the problems of interpreting results and scores obtained by them, since the teacher is not to use tests as ends in themselves, but as means to improving her methods of teaching. The tests are to furnish objective standards to be attained in the classroom by pointing out where over-emphasis must be avoided and where more emphasis is required. The introductory chapter demonstrates the inaccuracy of the prevailing marking system and states the advantages of objective methods of measuring the results of teaching. At the end of each chapter a number of stimulating questions and topics for further study are given. Many illustrations and tabulated results serve to make the discussions as concrete and helpful as possible. A valuable appendix gives detailed directions for ordering the right kind and amount of printed blanks and other test-material, with prices and sources. Everything is done to make the book practical and helpful to the elementary teacher.

L. R. G.

LEWIS M. TERMAN. *Condensed Guide for the Stanford Revision of the Binet-Simon Intelligence Tests.* Houghton Mifflin Company, Boston, 1920, pp. 32.

This "Guide" is based on Terman's Measurement of Intelligence and "a similar guide [that] had been prepared in the Office of the Surgeon-General for use in the army" with the assistance of Dr. J. W. Bridges and Major H. C. Bingham. In the "General Directions" certain injunctions are re-emphasized in the form of "ten commandments," amongst which the warning against "wholesale coaxing and cross-questioning" is especially urgent, as perhaps most frequently sinned against, particularly by teachers. In place of the "*Record Booklet*" there is now issued an "*Abbreviated Filing Record Card*" of 8½ x 11 inches, which will prove much more convenient and less cumbersome than the Booklet. It seems almost unnecessary to add that neither the *Condensed Guide* nor the *Abbreviated Filing Record Card* should be used by the novice in mental testing. L. R. G.

The following publications have been received:[1]

PHYLLIS BLANCHARD. *The Adolescent Girl.* Moffat, Yard and Company, New York, 1920, pp. 242.

JUNE E. DOWNEY. *Graphology and the Psychology of Handwriting.* Educ. Psychol. Monographs No. 24. Warwick & York, Baltimore, 1919, pp. 142.

HAROLD O. RUGG and JOHN R. CLARK. *Fundamentals of High School Mathematics.* World Book Co., Chicago, 1919, pp. 368.

DANIEL STARCH. *Educational Psychology.* The MacMillan Company, New York, 1919, pp. 473.

EDWARD K. STRONG, JR. *Introductory Psychology for Teachers.* Warwick and York, Baltimore, 1920, pp. 233.

M. R. TRABUE and F. P. STOCKBRIDGE. *Measure Your Mind.* Doubleday, Page and Company, New York, 1920, pp. 349.

---

[1] *Mention here does not preclude further comment.*

James R. Angell. *The Organization of Research.* Reprinted from *The Journal of Proceedings and Addresses of the Association of American Universities,* 21st Annual Conference, November, 1919, pp. 27-41.

W. V. Bingham. *The Division of Anthropology and Psychology of the National Research Council.* Reprinted from *Science,* N. S. Vol. LI, April 9, 1920, pp. 353-357.

Frank G. Bruner. *Sixty-fourth Annual Report of the Board of Education,* Chicago, for the year ending June 30, 1918, pp. 49-111.

Lotus D. Coffman. *Teacher Training Departments in Minnesota High Schools.* General Education Board, 1920, pp. 92.

Honorio F. Delgado. *La Psicologia de la Locura.* Reprinted from *El Siglo Medico,* Madrid, 1919, pp. 20.

——— ———. *El Psicoanalysis,* Lima, 1919, pp. 58.

Edward A. Pace and Others. *Psychological Studies from the Catholic University of America.* Psychological Monograph, Vol. xxvii, Nos. 118-122, 1919, pp. 515.

J. E. Wallace Wallin. *The Field of the Clinical Psychologist and the Kind of Training Needed by the Psychological Examiner.* Reprinted from *School and Society,* Vol. ix, No. 225, April 19, 1919, pp. 463-470.

——— ———. *The Achievement of Mental Defectives in Standardized Educational Tests.* Reprinted from *School and Society,* Vol. x, No. 244, Aug. 29, 1919, pp. 250-256.

The Educational Record. Vol. 1, January, 1920. Editor: Samuel Paul Capen. Published by the American Council on Education, Washington.

Journal of Educational Research. Vol. 1, No. 1, January, 1920. Editor: B. R. Buckingham, Univ. of Illinois. Published at Menasha, Wis., monthly, except July and August.

Philippine Journal of Education. Vol. 1, No. 7, April, 1919. Editor: Augustine S. Alonzo. Published by the College of Education, University of the Philippines, Manila. Monthly, except May and June.

Wisconsin's Educational Horizon. Vol. 2, Nos. 4-5, March-May, 1920. Editor: Edward A. Fitzpatrick. Issued bi-monthly by the Wisconsin State Board of Education, Madison, Wis.

Report of the Surgeon-General, U. S. Army, in 2 volumes, 1919. Government Printing Office, Washington.

War Department: Bulletin No. 11, March, 1919. Containing: *Physical Examination of the First Million Draft Recruits: Methods and Results.* Compiled under direction of the Surgeon-General, M. W. Ireland, by Albert G. Love and Charles B. Davenport.

Same authors prepared: *Defects Found in Drafted Men.* 66th Congress, 1st Session, *Senate Committee Print,* 1919.

THE
JOURNAL OF APPLIED PSYCHOLOGY

VOL. IV DECEMBER, 1920 No. 4

EQUALITY IN DIFFICULTY OF ALTERNATIVE INTELLIGENCE EXAMINATIONS.

By EDWARD L. THORNDIKE, Teachers' College, Columbia University

In investigations made to secure equal difficulty in the alternative forms of the Intelligence Examination arranged by the writer for use by Columbia College and other institutions, certain facts of general interest have been found, which I report at this time.

One element of the examination is a 12 page team of 13 tests, containing 144 tasks in all. There are 15 forms of this. They were made somewhat hastily from the harder tasks of some of the Army Alpha tests, and other tests known to correlate well with intelligence, extended by similar tasks improvised by Dr. A. L. Rogers, Dr A. I. Gates, the writer, and others. No great care was given to making the 15 corresponding tasks, or even the 15 corresponding tests, equal in difficulty. Only that was done that could be done by common-sense measures without using up much time or money in experimentation. There was in fact no time available in the case of the original use of these tests (the selection of officer material in the Air Service of the United States Army.) The degree to which the 15 forms are equal in difficulty thus measures the result that may be expected in 15 forms of a 30-minute examination of 144 items, arranged with common-sense precautions.

Ten of the fifteen forms are nearly enough equal for any purpose to which such examinations are likely to be put. A, E, J, L and N are approximately equal; C and I are about 1% easier; F, K, and M are about 1% harder; H is about 3% easier; D about 5% easier; G and O are about 2½% harder; B is about 5% harder. These estimates may be in error in respect to single forms, but the general drift is substantially correct. One of the quickest, and probably one of the cheapest, ways to get any given number of alternative examinations of approximately equal difficulty, is to make about 1½ times as many as will be needed, as nearly equal as common-sense procedures can make them, test all, and retain those which are up to specifications. The equality thus secured is a genuine equality of the different forms as they are actually used.

For most uses of these examinations only one of the fifteen need be discarded, since it is wise in general to use two forms together to make one hour of this sort of work. Seven such pairs differing as pairs by only 1% in difficulty can be had. These seven sets of two each thus furnish a convenient instrument for repeated measurements of the same individuals, or for the comparison of groups where there is danger that there may be gossip concerning the examination. The error of the instrument is inappreciable in comparison with the chance variation of individuals.

The evidence for the above conclusions is as follows:—

The material for measuring the differences in difficulty of the various forms consists chiefly of four sorts: (a) records of over 800 individuals taking two or three forms in succession without any fore-exercise; (b) records of over 500 individuals taking two or three forms in succession after fore-exercise; (c) records of about forty individuals each of whom did all fifteen of the forms, in the order A, B, C, D . . . . O.

First let us get certain very reliable differences estimated. Forms A and J being taken in that order, the median difference is $J-A = 12 \pm 0.7$ for 154 college students. Forms A and F being taken in that order, the median difference is $F-A = 10 \quad 0.7$ for 160 college students. So $J-F = 2 \pm 1.0$. Forms F and K being given in that order, the median difference is $K-F = 2\frac{1}{2} \pm 0.3$ for 351 college students. Forms J and K being given in that order, the median difference is $K-J = 3 \pm 0.5$ for 144 college students. So $J + 3 = F + 2\frac{1}{2}$ or $J-F = -0.5 \pm .06$. Combining all the results $J-F = .75 \pm .75$, or approximately $J = F + 1$. Forms B and F, and A and J, being given in those orders, the median difference $F-B = 15$ for 200 college students; the median difference $F-A = 10$ for 160 college students. Hence, we have $B + 15 = A + 9$ or $B = A-6$. Forms A and G being given in that order to 113 college students, we find the median difference to be $G-A = 9$. Since $F-A = 10$, $J-A = 12$, $G = J-3$ or $G = F-1$; but since, in the long run, $F = J-1$, we may put

$2G = J\text{-}3 + J\text{-}2$ or $2G = F + 1\text{-}3 + F\text{-}1$, whereby $G = J - 2\frac{1}{2}$ or $G = F - 1\frac{1}{2}$.

The forms Practice form, B, and C being given in that order, we have a median difference $C - B = 12$ for 86 individuals. The forms Practice form, B, and F being given in that order, we have a median difference $F - B = 10$ for 126 individuals. Hence $C - F = 2$ or $F = C - 2$.

So far we have $J = F + 1 = G + 2\frac{1}{2} = C - 1$, and $B = A - 6$.

Assuming that the general practice gain from second trial to third trial is 3,[1] we can evaluate H and K as follows:—

In X G H, the median $H - G = 9$ for 113 individuals. $H - G$ irrespective of position would then be 6, and $H = G + 6$. Since $J = G + 2\frac{1}{2}$, $J = H - 6 + 2\frac{1}{2}$. So $H = J + 3\frac{1}{2}$. In X F K, the median $K - F$ for 151 individuals is $1\frac{1}{2}$. $K - F$ irrespective of position would then be $-1\frac{1}{2}$, and $K = F - 1\frac{1}{2}$ In X J K the median $K - J$ for 144 individuals is $3\frac{1}{2}$. $K - J$ irrespective of position would then be $\frac{1}{2}$ and $K = J + \frac{1}{2}$. Combining the two lines of evidence on K, $2K = F + J - 1$, whereby $2K = F + F + 1 - 1$, and $K = F$, or $2K = J - 1 + J - 1$, and $K = KJ - 1$.

Assuming 12 as the practice effect of second over first position when no fore-exercise is given, we can evaluate A. We have (from the sequences A F, A G and A J,) median differences for 160, 113 and 156 individuals, as follows: $F - A = 10$, $G - A = 9$, $J - A = 12$, including the practice effect. Excluding it, we have: $F - A = -2$, $G - A = -3$, $J - A = 0$. From these $3A = 3J + 1\frac{1}{2}$.

So far we have $J = F + 1 = G + 2\frac{1}{2} = C - 1 =$
$H - 3\frac{1}{2} = K + 1 = A - \frac{1}{2} = B + 5\frac{1}{2}$.

Using these facts we may estimate the amount of practice effect to be allowed for in those who took all fifteen forms in the order A, B, C, . . . O. If there is no practice gain from F on, $J + K$ will equal $F + G + 2\frac{1}{2}$. In fact $J + K = F + G + 4\frac{1}{2}$ in the group of 38 individuals who took all forms. There seems thus to be a gain of about $\frac{1}{2}$ point per form due to practice at this stage. If there is no practice gain from C to F G, $F + G$ will equal $2C - 5\frac{1}{2}$. In fact $F + G = 2C + 10$. So the practice seems to be $+ 7\frac{1}{2}$ from G to F G, or about 2 units per form.

---

[1] 3 is the median of 13 such cases. Also, G and J being so taken by 26 individuals, the median $J - G$ was 5. Since $J = G + 2\frac{1}{2}$, the practice effect would be $2\frac{1}{2}$.

We may then assume that if all the forms were of approximately equal difficulty, the improvement from C to O would be approximately $2\ 2\ 2\ 1\ \frac{1}{2}\ \frac{1}{2}\ \frac{1}{4}\ \frac{1}{4}\ \frac{1}{4}\ \frac{1}{8}\ \frac{1}{8}\ \frac{1}{8}$. The average scores for C, D, E, etc. for the 239 individuals tested were, in order, 90, 97, 93, 97, $92\frac{1}{2}$, $96\frac{1}{2}$, 99, 99, 95, $97\frac{1}{2}$, 97, 99, 97.

The amount by which each form is easier or harder than the average of all the forms may then be estimated approximately as:
$C = Av + 1$, $D = Av + 6$, $E = Av$, $F = Av + 3$,
$G = Av - 2\frac{1}{2}$, $L = Av + \frac{1}{2}$, $I = Av + 2$, $J = Av + 1\frac{1}{2}$,
$K = Av - 3$, $L = Av - \frac{1}{2}$, $M = Av - 2$, $N = Av$, and
$O = Av - 2$.

Using these facts to extend and amend the values already obtained, we have, in relation to the average:—
$A = +\frac{1}{2}$ $B = -5\frac{1}{2}$, $C = +1$, $D = +6$, $E = 0$,
$F = -1$, $G = -2\frac{1}{2}$, $H = +3$, $I = +2$, $J = 0$,
$K = -1$, $L = -\frac{1}{2}$, $M = -2$, $B = 0$, $O = -2$.

We may now check these estimates by estimates derived from entirely different data.

Assuming that the practice effect of second over first position when fore-exercise precedes both[2] is 7, we have $G - B = 12 -$ practice and $C - B = 10 -$ practice from 126 and 86 individuals respectively. Hence $B = F - 3$ and $C = B + 5$. Since F is accurately determined as $Av - 1$, we have $B = Av - 4$ and $C = Av + 1$.

313 individuals having Practice form, D, and E, showed a median $E - D = 4\frac{1}{2}$. Apart from practice, then, $E - D = -2\frac{1}{2}$. 28 individuals having Practice form, G, and H, showed a median $H - G = 9$. Apart from practice, then, $H = G + 2$. 20 individuals taking O and E after fore-exercise, showed a median difference of $E - O = 10$. Allowing for practice $E - O = 3$. $O = E - 3$.

15 individuals did G, H, and I, five in the order G, H, I, five in the order H, I, G, and five in the order I, G, H. The median scores were 108, 115, and 110, whence $H = G + 7$, $I = G + 2$, $H = I + 5$.

27 individuals did forms F and G in that order without fore-exercise. The median $G - F$ was 10. Allowing 12 for practice, G is 2 harder than F, or $G = F - 2$. 7 individuals did forms D and J in that order, without fore-exercise. The median $J - D$ was 7. Allowing 12 for practice, J is 5 harder

---

[2] This assumption is made on the basis of the fact that, including the effect of practice, $C - B = 12$ for $n = 86$, $F - B = 10$ for $n = 126$, $E - D = 4\frac{1}{2}$ for $n = 313$, $H - G = 9$ for $n = 28$.

than D, or D = J + 5. 36 individuals did forms F and J in that order without fore-exercise. The median J — F was 13½. Allowing 12 for practice, J = F + 1½. 33 individuals did forms G and J in that order without fore-exercise. The median H — G was 15. Allowing 12 for practice, J = G + 3. 26 individuals did forms G and J in that order as second and third. The median difference was J — G = 5. Allowing 3 for practice, J — G = 2, or J = G + 2.

28 individuals did forms J and L in that order without fore-exercise. The median L — J was 11. Allowing 12 for practice L = J — 1. 15 individuals did forms G and K in that order as second and third. The median difference was K — G = 4. Allowing 3 for practice, K — G = 1 or K = G + 1. 36 individuals did H L in that order as second and third. The median difference was L — H = — 1. Allowing 3 for practice, L — H = — 4 or L = H — 4.

Similar data for the sequences K M, J M and L M give the following:

n = 9, median difference M-K = O. So M — K = — 3; M = K — 3.

n = 40, median difference M-J = 2½. So M-J = — ½; M = J — ½.

n = 28, median difference M-L = 4½. So M-L = — 1½; M = L + 1½.

Similar data for the sequences C N, K N, and L N give the following:

n = 9, median difference = O. N — C = — 3.
n = 7, median difference = 6. N — K = 3.
N = 4, median difference = 1. N — L = — 2.

Using these facts we have the following estimates to compare with those previously obtained.

		Previous Estimate.	Estimate from Accessory Data.
A	Av	+ ½	
B	"	—5½	—4
C	"	+1	+1
D	"	+6	+4
E	"	0	
F	"	—1	
G	"	—2½	—2½
H	"	+3	+2
I	"	+2	- ½
J	"	0	+ ½
K	"	—1	
L	"	— ½	—1
M	"	—2	—1
N	"	0	—1
O	"	—2	—3

On the whole we shall not be in serious error if we take A = E = J = L = N, counting C and I as 1 point easier, F, K and M as one point harder, H as 3 points easier, D as 5 points easier, G and O as 2½ points harder, and B as 5 points harder. The probable error of these estimates ranges from about ¼ to 1½ points.

CORRELATION OF ARMY ALPHA INTELLIGENCE TEST WITH ACADEMIC GRADES IN HIGH SCHOOLS AND MILITARY ACADEMIES

By H. E. Burtt and G. F. Arps, Ohio State University

Introduction

Numerous studies have been made in recent years comparing mental tests, and especially intelligence tests with academic grades. The underlying hope seems to be that a student with good mental ability or intelligence (original nature) will do well in school or college and receive good marks and that, hence, it will be possible to predict future academic status on the basis of tests which measure intelligence. One of the most notable results of practically all such studies is the low correlation obtained. A correlation of .50 seems to be the exception rather than the rule, and prediction based on a correlation much smaller than this is of doubtful value in the individual case.

There are at least two errors which might contribute to these low correlations: either the tests are not an accurate measure of mental ability or the school marks are not an accurate measure of academic ability. Either measure is of course vitiated if the work is not performed with maximum incentive. It is probable that most pupils work with a fair approach to their maximum ability in taking mental tests. The test is usually brief and does not reach the point of producing "annoyingness" as is often done by school work. Furthermore the pupil being unaware of the real purpose of the tests assumes that a good deal is at stake and, consequently, does his best. It is probable on the other hand that many students do *not* work with maximum incentive in their studies or do not devote to them sufficient time to manifest their maximum ability. Any teacher would heartily endorse this statement. The writers feel that this latter alternative is a possible explanation of the low correlations generally found between intelligence tests and academic status. Suggestions of this sort have been made before.[1] It would seem further that certain types of instruction hold students more closely than other types to their maximum intellectual ability. One of the writers found a noticeably closer relation between the adult "point scale" (Yerkes) and grades of students in a psychology course which had frequent and vigorous written examinations than between point scale and total academic marks for a similar group.[2]

---

[1] Anderson, J. E., Intelligence Tests of Yale Freshmen, *School and Society*, Vol. XI, 417-420.

During the summer of 1918 one of the writers was in charge of the mental examination of a group of R. O. T. C. men at Camp Custer, Michigan. Most of the members of the group were at the time in the midst of their course in high schools or military academies. These two types of institution differ considerably in methods of instruction especially with respect to supervision of study, general supervision of pupils and systems of demerit (punishment). It seemed possible by comparing the intelligence tests and academic marks for those groups to determine whether the methods of instruction do differ in the extent to which they elicit maximum intellectual performance in school.

Method

Academic marks prior to army service were obtained from the heads of schools and academies which these R. O. T. C. men had previously attended. The information was obtained on a blank calling for the average mark of a student in science, mathematics, languages (including English), history and "other subjects" and also the number of courses of each sort that were taken. Most of the marks obtained were in numerical form so that statistical treatment was possible. It was an easy matter to obtain a single average for total academic work by weighting each branch of study in proportion to the number of courses taken in that branch.

It is obvious that marks obtained from different schools with their different standards of marking could not be validly grouped together in crude form to obtain a correlation coefficient. Even if the marks were thrown in to terms of average or passing mark there would still be a source of error in the scatter for different schools. Consequently, each school or academy was taken separately and the mean and standard deviation obtained for total academic marks of the R. O. T. C. men of that school and the mean and standard deviation found also for the various school branches for those same men. Each individual mark was then put in the form of X divided by standard deviation where X is deviation from the mean. Thus a student with a score of + 1.5 in school A is of the same academic ability as a student with a score of + 1.5 in school B, although the passing mark and the scatter of the students in the two schools are entirely different. The score simply indicates what proportion of the standard deviation for this school the person is above or below the average for this school. Scores of this sort were then available for all the features of interest, and it was valid to include all the military students in one group and the high school students in another. No school was included from which there were fewer than ten R. O. T. C. men. These scores were then plotted against the Army Alpha Intelligence Test and the correlation computed by the products—moments method.

---

[2] Yerkes, R. M. and Burtt, H. E. The Relation of Point Scale Measurements of Intelligence to Educational Performance in College Students. *School and Society*, Vol. V, 535-540.

Results

The correlations are based on 118 men in five military academies and 93 men in eight high schools. The groups themselves are very similar in intelligence as shown by the test. The military academy group has a mean score of 129 and a standard deviation of 25.6, while the high school group has a mean score of 130 and a standard deviation of 25.8. The principal interest lies in the comparative correlations for the two groups. These are given in the tables.

TABLE I

Correlation Alpha Total With:	*Military Academies*	*High Schools*
Academic total	.39	.19
Mathematics	.32	.16
Science	.47	.18
Language	.42	.22
History	.40	.25
P. E. (Approximate)	.05	.07

TABLE II

Correlation Academic Total With:	Military Academies	High Schools
Test 1	.41	.06
Test 2	.36	.14
Test 3	.19	.27
Test 4	.41	.27
Test 5	.32	.26
Test 6	.25	.16
Test 7	.45	.18
Test 8	.25	.37
P. E. (Approximate)	.06	.08

Table 1 gives the correlation of the total Alpha score, that is the general intelligence rating, with the total academic marks and also with marks in several school branches. The correlations are of about the same order that is usually found in similar studies. The most noticeable thing in the table is the correlation of intelligence and total academic ability of .39 for the military group as contrasted with .19 for the high school group. This is a most significant result of the study. The writers feel that it indicates a superiority of methods of instruction in military academies from the stand point of holding the students to their maximum mental efficiency. This result does not seem to be due primarily to any single school subject. The difference between the correlation for the two groups holds rather consistently with various school subjects as indicated in the other figures of the table. The difference seems to be slightly greater with science and slightly less with history but it is very doubtful if this fact is of significance. It is impossible that the discrepancy between the two groups could be due to a greater homogeneity of the high school group in intelligence because the standard deviations show that the distribution curves of the two groups are practically identical.

Table 2 gives the correlation of total academic marks with the eight different tests of the Alpha Examination. This table is not as significant as the first. There are two tests, 3 (practical judgment) and 8 (information) in which the correlations are slightly higher for the high school group but the differences are not significant when the P.E. is considered. The greatest discrepancies in the other direction are in test I (directions) and test 7 (analogies). This would suggest that these latter tests are somewhat more indicative of the instruction differential between the two types of school.

It seemed worth while to investigate the linearity of the regression in the case of the two principal correlations of the study. The correlation ratio (eta) was found to be .52 for academic total on Alpha total in the military group and .49 in the high school. The corresponding correlation coefficients were .39 and .19. Blakeman[3] suggests several formulae which give the ratio of the difference between eta and r, or of the difference between their squares to the probable error. His simpler formula (recommended for rough work) gives a value of 2.8 for the military regression and 3.2 for the high school. His more exhaustive formulae, (one giving ratio of difference of squares to probable error and the other giving ratio of difference of first powers to probable error) agree closely in yielding values of 3.0 and 3.9 for the respective groups. Blakeman considers that this value should be less than 2.5 (corresponding to a probability of 9 to 1) to indicate linearity.

---

[3] Blakeman, J. On Tests for Linearity of Regression in Frequency Distributions. *Biometrika,* Vol. 4, 332-350.

These figures indicate that for the high school regression at least the observed values fit some curve more closely than they fit a straight line. Inspection of the scatter plot suggests a considerable concentration in the quadrant for high intelligence and low academic grades and the possibility of the regression line curving somewhat in to that quadrant. Many of the individuals in this quadrant are doubtless those who have good ability but are not induced to use it in school. This feature is not nearly as clear in an inspection of the plot for military students and furthermore Blakeman's test shows this regression to be more nearly linear. These facts are a further corroboration of the main thesis of this study.

Conclusion

The low correlation generally found between intelligence tests and academic marks may be to a considerable extent due to the fact that methods of school instruction do not hold students to their maximum efficiency. The correlation of Army Alpha with academic marks in high schools and military academies yields an appreciably higher correlation for the latter group. This is not caused by differences in homogenity of the groups nor does not seem due primarily to any special school subject but to be a general characteristic throughout the curriculum. This suggests that methods of instruction in the military academy, such as general supervision, supervised study and a system of reward and punishment, are more apt to hold the student to his maximum intellectual ability.

SHOULD PSYCHOLOGY BAKE BREAD*

ELIOTT FROST, Rochester, N. Y.,

I

Baking bread is a human task, popular rather generally, and over a sufficient number of years to give it standing among the very best people. Perhaps for these several reasons we allow it in our thinking to typify the practical. Should psychology in this sense bake bread, and if so what kind of bread should it bake?

It is to be generally admitted that a science implies the observation and collection of raw facts, and the correlation of these facts under laws. Such laws once determined may, if we so desire, then be utilized to secure ends—either theoretical purposes which shall lead us on to additional facts; or practical ones which shall help us to justify our studies by enabling us to aid in the solution of those very human problems that we all, from time immemorial, have faced.

There is no one of scientific temper, who would care to defend the thesis that truth is " its own excuse for being " if by the statement we mean that the question—" What is this true thing *good for*," is a wholly irrelevant one. Whatever our personal inclination may be as to the propriety of our making applications of our data once they are assembled; whatever our belief as to the timeliness of attempting application of a particular science, or of attempting it under any special conditions, there is a very general admission that both the excuse and the incentive for research, lies in the fact that the results of it may, in due and proper time and under some auspices, receive justification through use. Such employment of our results may be, and may properly be, indefinitely postponed or even left for future generations to harvest in benefits, but in our minds there must always be implicit expectation that somebody sometime will somehow translate into terms of life those formulae which we here and now devise and amplify. Otherwise we are lending ourselves, our energy, and our training to an avocational pursuit which perhaps may be permitted as temporary recreation but for which we have no right to receive the support and rewards of society. Thus to employ our entire time is to be as truly a parasite upon the social host as if our days were successively spent on the golf-course or at the chess-board.

---

*Read before the local branch of the Academy of Science, Rochester, N. Y., March 1920.

One of our psychologists is alleged to have said—" I should be just as much interested, were I a geographer, in plotting some insignificant South Sea Island of no consequence to anyone, as I would in plotting the Island of Manhattan." Such a statement, if true, needs interpretation. If by that he means that he, personally, would not be interested in accomplishing practical results, and seeks only for truth (which someone else of different temperament might subsequently apply)—that is one thing. On the other-hand, if he means that he would willingly devote his time to the investigation of a pursuit which would have and could have no advantage for anyone—that is an entirely different affair. Truth for its own sake is an excellent slogan as far as it goes but it never can or should tell the whole story.

Let us therefore analyze our original question with care, distinguishing it from the entirely different question—" Should *Science* Bake Bread? " In the latter question we leave out of account the age of our particular science and its particular development. In short, it may well be that we subscribe fully to the general applicability of science and still deny categorically that psychology is ready for such application.

II

At least two considerations may militate against the wise application of a science: First, the science may be so new and its conclusions in consequence so half-baked, that the premature utilization of them are of no value; or second, we run a liability of prejudicing the future of the science itself, not only in the minds of practical men but in the minds of our own investigators, so that their subsequent researches are partly, even though sub-consciously controlled, now no longer by a search for truth but by such biases and opinions as would be dictated by needs. This latter result, should it occur, would be not only unfortunate, but calamitous, for it would poison our springs at their source.

We are now, if this be granted, in a position to clarify and simplify our topic although we state it at greater length: " Is psychology, as such, already sufficiently developed so that it may hope, at least in part and wisely, to offer its formulae and conclusions as practical recipes toward the solution of human problems? "

It might seem, in the light of what has transpired, particularly during the last four years, that the question is being asked a decade too late, were it not for the fact that a considerable number of eminent American psychologists have persistently given a negative and *a priori* answer to the question; while many others, if of lesser note, after viewing the attempt of psychology to sell its services generally have been won over to a similar, though *a posteriori*, conclusion. Some in short have said " No, psychology should not attempt to apply itself " while others have added—" Such attempts as it has made have not justified themselves."

To those of you who have followed psychological discussions and psychological literature for the last fifteen years, from the time when Münsterberg railed against the application of psychology; through the period when he wrote his numerous books suggesting just such application; down to the present time when our psychological meetings have veritably centered about applied interests, it will seem that most psychologists have answered our query in the affirmative. Psychological applications to education, to medicine, to advertising, to personnel, and to industry have of late been presented to us in a bewildering swarm.

What is the net result first upon the object sought, and second upon the science which has offered its service? As to the first of these questions, namely, whether or not the results obtained have been satisfactory, it is the opinion of the writer that they have not, if we measure those results in terms of energy expended. Probably the application of psychology to medicine even in its indirect consequences comes nearer to offering results of which we may be proud than does any other field. Even here however we have had to contend with psychoanalytic orgies, psychotherapeutic buncombe and a congeries of lesser quackeries which have so enveloped the kernel of truth contained, as to handicap and postpone the success that might have come from a more conservative investigation of the premises upon which our somewhat hasty conclusions have been based.

On the other-hand, what we have failed to gain in therapeutic diagnoses and prognoses; in pedagogic recipes or industrial formulae, is made up for—and probably more than made up for—in other ways. We have brought the terms and the processes of psychology before a very wide and very important public. Since a few years ago when Mr. Wilson declared the trouble of business to be " psychological," men have ceased to look sheepish when they use the term. To-day we find the literature of industry naturally adopting psychology as a necessary, explanatory science.

Psychology suffers and always must suffer from a matter-of-fact point of view, by the circumstance that she is dealing with processes rather than masses and molecules. This aggravates the difficulty of our descriptions and explanations but in no way lessens the importance of them. On the contrary, the duty of spreading knowledge concerning human instincts, feelings, emotions, and reactions, was never so universally recognized as to-day. The fields are plowed and fallowed as never before for every constructive suggestion that psychology has to offer. For this happy result the embryonic essays of the last decade in psychology are chiefly responsible, however insignificant the specific success of any particular planting.

Those who are familiar with the work of Veblen's disciples: Helen Marot's " The Creative Impulse in Industry " and Ordway Tead's " Instincts in Industry;" those who have read the life of Carleton H. Parker or who know the work done by Mr. Robert M. Wolfe in his pulp and paper factories, those who have followed the development of Industrial Relations men, the Labor Manager, and the Employment Manager; those who have been studying the newer ideas of plant management, plant sanitation, bonus and profit sharing systems, and the training of foremen; those who have followed the course of labor adjustments in England through the Whitley reports and elsewhere; and those who are familiar with the innumerable boards, bureaus, commissions, and societies designed to ameliorate industrialism, and put back the loyalty that comes from content and creative opportunity—to all such one need not emphasize that the principles of psychology are being put to work in industry to-day, and as a direct result, undoubtedly, of what psychology has done in other fields already now for many years.

Quantitative changes in time always involve qualitative changes. In theory the ocean is but the aggregation of drops of brine. This ocean however knows of things of which the droplet never dreamed—winds and storms and waves, icebergs and shimmering sea, depths and foam, climate and colors and temperatures—all these things and many more. Just so has our industrial life grown from the small shop with the single employer and his handful of friendly assistants to a huge depersonalized corporation controlled by absentee capital and engulfing the human individual unit as one drop by the sea.

Ancient contacts are rendered impossible; specialization has demanded mechanization, the human has become the machine. Mass psychology has supplanted the zests and rewards of individual initiative. Our industrial patient has become ill and we are collectively concerned to discover some remedy for that which we could not apparently foresee or prevent. The human factor will no longer be denied. The drop refuses longer to be merely a part of its environment.

Either we must break up our great industries and return as some firms are doing, to a group of relatively small plants, distributed in relatively small communities (while maintaining perhaps a centralized control), return that is, to conditions where the individual worker may still maintain himself as an individual, and enjoy his home, his family and avocations; or, where the plant is necessarily large we must pyramid friendship so to speak, provide, that is, for open channels and organized contacts between management in the office and loyalty in the shop. The one thing we can no longer do is to ignore the vital human relationships which separate the independent spirit from the spirit of the slave.

III

We come then to the second portion of our problem. What is the effect of success upon the science itself and those who carry its torch?

One must confess the answer to this question is difficult,—not that there is so little to say but so much. If one felt that the achievements or even the program of theoretical psychology ten or fifteen years ago, (prior to the recent expansions into the domain of applied psychology) were dealing with important and illuminating issues then he would incline to the view that we were fallen into unprofitable times. On the contrary, if you will take pains to review the topics presented for consideration and discussion at the annual psychological meetings, say fifteen years ago, you will be impressed in the light of, and perhaps because of, subsequent and larger vision, with the sterility of most of the investigations undertaken. The incentives which have come to our science along practical lines, have not served as a healthful stimulus to more adequate research even for purely scientific interests. Just as the science of medicine has gone hand in hand with the practice of it, so it is possible and perhaps likely that within limits, psychology henceforth may fare. The danger is however always and near at hand that either because of public clamor or financial reward we shall turn too quickly from our laboratory into the more remunerative fields. The ideal is, that there should always be a number of earnest men fitted by temperament and training to carry on pure research. If this group is guaranteed, we can without trepidation permit another group to test out in the realm of practical affairs such tentative conclusions as our laboratory offers. If these two groups are recognized as distinct and so maintained, we should expect progress, both for our sciences and in the solution of their practical problems.

IV

Admitting, then, that bread should be baked and that the psychologist, under certain limitations may properly be a baker, brings us to the very concrete problem as to what kind of bread. The answer is, to be sure, somewhat speculative and more than commonly so just at present. The industrial bakeshop is stirred mightily and there is much rattling of pans and kettles. New cooks with favorite recipes run about. Has the psychologist a place among them?

Three distinctly wholesale tendencies, which are working toward industrial stability, are *first,* the increasing willingness of competitive employers to meet and share their problems and solutions; *second,* the fuller recognition of the shop foreman in the industrial plant to-day as a possible salesman of the management; and *third,* the growth of departments of Industrial Relations.

First, then, the increased willingness of competitive employers to meet and share their problems and solutions. The intensely competitive era is passing. Unprecedented problems on the one hand, and on the other, the co-operations of the war to make the Liberty motor, to fabricate ships, and to place invention or patent at the Government's disposal, has hastened a tendency already apparent at the beginning of the war toward combinations and co-operations not only of captial but of idea. A single illustration of such development is the Industrial Management Council of Rochester, N. Y. now entering upon its fourth year of service, among forty odd of the leading manufacturers of that city. This Council is subdivided into seven groups: General Managers, Office Managers, Production Methods, Employment and Service, Superintendents, Cost Accountants, and Taxation. Its primary purpose is to swap methods and experience, and to gain additional information through surveys, investigation, and reports, as to the best manufacturing practices from concerns inside and outside of the city.

The Council is operated by a Director and an Executive Committee, in addition to a chairman and program committee of each group. During the year, with the exception of the summer, the groups meet bi-weekly, and are addressed by experts whose talks then form the basis of questions and discussion. So far as is known, this program is unique and not duplicated in any other city of the country, though many cities have indicated their intention of forming similar organizations.

Out of the multitude of alleged panacea offered to the harrassed manufacturer to-day, one note is permanent and dominant: The necessity for more genuine mutual understanding and a more practical co-operation. This need and this principle hold true, not only as between the management and the worker, but among manufacturers and executives in the interest both of themselves and the worker. In the past, there has been the problem of the closed shop—the shop, that is, closed to the rival manufacturer and often in consequence closed to new ideas.

Manufacturers are now increasingly coming to feel that the welfare of any industry vitally depends upon the welfare of all industries, and that through the interchange of ideas, more will be gained in the long run. This would be offset through the temporary loss of advantage from passing along a good idea to one's neighbor. Upon this conviction the Industrial Management Council is built, and in conformity with this thesis, it operates.

All member concerns of the Industrial Management Council are represented in the Rochester Chamber of Commerce which houses the central offices of the Council and provides a meeting place for its members. Group meetings occur on a regular schedule throughout the year, are planned in advance, and are usually preceded by an informal dinner. The Chamber furnishes rent, heat and light; other expenses of the Council are defrayed by assessment upon each firm and according to a definite formula. Occasional group meetings of foremen, upwards of 1,000 at a time, are held under the auspices of one or more of the groups, and either foremen or workmen at the bench are always welcome as guests of members, since it is the interest of the Council to safeguard and promote the success and welfare of the worker no less than that of the management. Some synthesis must be found as between the thesis of capital and the antithesis of labor. Their goal is, ultimately, the same, and we can no longer afford to postpone its realization by quarreling over methods. The Industrial Management Council is a sober and honest attempt to sit at a common council table, to view industry as a whole, and thereby to come to some conclusion which shall benefit all partners to it.

A *second* hopeful sign of our industrial times is the fuller recognition of the shop foremen in the industrial plant to-day—the recognition of the need not only for choosing the right kind of man but for giving him the right kind of education once he is found. It is coming to be increasingly true that the foreman marks the measure of success attained by the industry, and makes more than a nominal contribution to it. He is in a position to sell the ideas of the management to the employee and to sell the employee to the management. He is the lock in a canal. If he fails to work, there is danger that the waters on either side of him may grow stagnant and breed discontent.

Nor are factors lacking to produce just such a situation. In the first place the foreman has a bad history to overcome. He has been treated in times past as a cog in a machine, whose primary duty was to get out production at any cost. Physical prowess, a bull-dozing manner, the inspiring of fear;—these formerly were considered assets and men were picked as foremen because of their possession of them. The modern tendency in industry to adopt methods of suasion rather than compulsion has left us with this body of men trained to the exercises of authorities which we now deny them. If our newer ideals are to persist, we must either select new foremen or re-educate such as we have.

Against the effectiveness of this key-man in industry, militates also the vivid radicalism of a relatively small group of agitators. These men speak with a sincerity, or an apparent sincerity, worthy of a better cause. If we could but grant the major premises, we must admire their minor premises and their conclusions. They use short Anglo-Saxon terms, they speak and write so that even the simple-minded can understand. There are said to be nearly 1,200 periodicals, regularly published, in this country, sufficiently red of hue as to come under the category of 'Bolshevik.' Such husky germs can be killed by one specific only: the truth. The foreman is the one man who can quickly, by word and deed, morning, noon and night, administer this remedy. But he must first possess it. That is the responsibility of management: to give him the facts and keep him sold, lest, instead of becoming a doctor for industrial discontent, he himself catches the disease and spreads it.

Every foreman can become a salesman of the management. Few workmen know where the contents of their pay envelope comes from. Few know what 'overhead' means. Few, if any, have an idea of the cost of doing business or the risks involved. Yet the average worker is interested in economic education, although he may not know it by that name. Part of this interest has been aroused by this extraordinarily clever literature, mostly radical, which would be very important if true, but which is not true. If a man is told that his employer is making ten or even fifty dollars a day from his labor; if he is told that, day after day by word of mouth and printed paragraphs, and no one takes the trouble to tell him it is not so, he is likely in time to believe it. I would and you would. The large employer of labor cannot go into his factory and chat daily with all of his workers. He can however give his policies and his ideals to the chiefs under him, and see in turn that they are passed down the line to and through the foreman. Just as the Sergeant in the Army reflects his Colonel, so does the foreman in the shop reflect his chief in the office. If management does not take the pains to instruct its foremen, let it not expect that these men will keep the channels open, to and fro, between office and bench.

There are many present-day attempts to educate the foreman by courses of intensive study and lectures, notable among them the work of the corporation schools of New York, the National City Bureau, and the National Speakers Bureau. Already however, these admirable educational programs have met with criticism. In some cases it has been felt that the foreman was being educated beyond his intelligence and into a restlessness toward, rather than a loyalty to, his job. It is claimed further that the teaching is valuable, as most teaching largely must be, only in the hands of excellent teachers, and such personnel is difficult to secure and repay. Finally the economic laws which govern our industrial life are operating also here and foreman training is likely to be advocated in direct proportion as the supply of foremen is limited. In other words when labor is plentiful the average manufacturer will simply select the best foremen and bother himself very little about the remainder.

A *third* index of a fortunate industrial evolution is the increasingly significant rôle played by the departments of Industrial Relations, and the correspondingly larger number of men with special training and ability who seek them. Leaders in this new enterprise are receiving large salaries, and, speaking generally, they earn them.

During the war, the Army and Navy each added a Morale Branch to its General Staff; too tardily, it is true, to achieve the effective success that an earlier recognition of the principles would have permitted, but sufficient to sell the idea to the permanent and regular establishment. Since the war, with a demand for production still abnormal, under-production has been a common lament, and is correlated with high labor turnover, strikes, and mental hysterias beyond precedent. William James long ago pointed out that none of us is maximumally efficient; that we possess latent sources of energy we seldom, if ever, tap. Added to these inefficiencies we now have those which come from a sullen and unwilling spirit, from a moral refractory period following the intensities of the war, and from the aggravating but separate circumstance that labor and capital have lost the personal touch.

Industrial relations are passing through a critical phase. Viewed by many manufacturers as a necessary sop, grudgingly granted, there is a tendency to take the short, rather than the long look, in a hopeless attempt to return to things as they were. On the other-hand, manufacturers with wider vision who are less interested in concerns of the moment and more interested in building for the future, contemplate this newer machinery which the times have set up, as an asset whose interest return will more than justify capital expense. If industrial safety, industrial housing, hiring and transfer of employees, training of workers, mental incentives and the sharing of responsibility, are looked upon as beneficences, the first easement of the labor market will sweep them into discard as economic irrelevancies. But the thoughtful manufacturer no longer so considers them. He is willing to buy and pay for loyalty no less than brains and muscle; he realizes that he must attract virtue no less than reward it.

Just how many manufacturers will fall into this category of the 'thoughtful' when times change, and the man seeks the job, rather than as at present, when the job seeks the man, it is difficult to predict. Until it can be established beyond peradventure that it pays to educate the man *into* his job and *in* his job, gratuitous education is not kindly taken to. One's prediction depends in part upon his interpretation of human nature. Either for lack of vision, or for lack of money, or both, most industries will probably make shift to use the human material that they have on hand in much the same condition as they find it. For those of us who are interested in the dissemination of education, this is discouraging but we must take the world as it is. It does not help much to criticise life. We can modify its formulae but slowly. Because we cannot secure high-minded unselfishness on the part of all of our employers is no adequate ground for pessimism. Enlightened self-interest is quite a different affair from selfishness. If we can prove that the goal we set in our co-operative policies, our foremens' training, and our industrial relations, ministers to the good of the individual by ministering to the good of all, we shall have taken a step forward and up.

Since most industrial problems are at bottom due to ignorance, they can be satisfied only by educational programs of some sort. There are many signs of a closer liaison between our universities and our industries. Several of our colleges and technical schools, notably the University of Cincinnati, have formed a close alliance with their neighboring industries, alike to the benefit of the factory and of the school. Some of our industries too, such as the Goodyear Tire Company, the Ford Motor Company and others, have established universities within their own plants, and financed them with a generosity which compares not unfavorably with many of our smaller colleges.

V

Specifically, the technique of the psychological laboratory, is not yet sufficiently standardized to translate it effectively to industry. Such work as that of Link is a good step forward but it is necessarily limited in scope. The major premise that the human problems of industry are psychological is well established and curiously enough, industry and not psychology has been the first to recognize it. But the minor premise of application still is lacking and likely to remain so until we develop a larger group of men whose interests and background are both psychological and industrial. The man without psychological training will miss the forest for the trees; the man without industrial experience will see the forest but miss the trees. All of which is only to say that industry inclines to be over-practical, measuring results too objectively and quickly in profit and loss; while the laboratory psychologist leaves out of his reckoning a hundred and one variables due to type of product, nationality of worker, history of management, and the factors of production control, which vary with each industry.

Industry to-day is needing the psychologist, certainly, or at least the man of psychological temper; but he must come to her with his overalls on; must supplement the truth he knows with facts gained by experience, and must be as eager to learn as to teach.

Underdone bread is sometimes worse than no bread, and makes for indigestion. If the psychologist is to do any industrial baking, he must learn first to cook, for as everyone knows, cooking is an art and its secrets lie outside of books.

A FURTHER DEVELOPMENT OF EMPLOYMENT PSYCHOLOGY

By Henry C. Link

Aside from the mere selective and classifying value of tests there are connected phases of employment procedure in which the psychological technique may be very valuable. The psychologist in industry finds in his technique, an instrument which sometimes gives him a great advantage over the ordinary employment man when it comes to analyzing labor supply, adjusting rates, discovering inconsistencies in existing employment procedure, etc., etc.

The following study, made in an industrial concern in which tests were a supplementary part of the employment procedure, indicates several directions in which the industrial psychologist may look for practical results.

During a period of six months, the psychological section of the personnel department had given clerical tests to 616 women who applied for clerical work. The method by which these tests were put into operation and the results followed up is described elsewhere. At the time when the study described here was made, the success of the tests had already been established.

Of the 616 applicants examined—

287 recommended by the psychological section had been hired;

137 had been recommended but not hired;

173 had not been recommended and not hired;

19 had not been recommended but hired nevertheless.

The employment office had therefore followed in the main the recommendations made by the psychological section. However, the large number of applicants recommended but not employed, taken in conjunction with a constant shortage of clerks, made the following study advisable. The purposes of this study were:

1. To open up new sources of clerical workers.
2. To determine on the basis of tests which had already proved satisfactory, how the performance of the applicants examined in the employment office compared with the performance of potential candidates who had not yet applied for positions.

3. To determine a basis for adjusting earning rates so as to make possible the selection of the most promising potential applicants.

4. To point out discrepancies which were known to be present in the existing employment procedure.

A series of group tests, therefore, was given to (a) the senior class of the commercial high school, 181 pupils; (b) the beginners and advanced class of the X business college, 42 and 31 pupils respectively; (c) the advanced class of the Y business college, 38 pupils. The principals of these schools readily fell in with the plan of the study when its purpose was explained to them. The tests given were: (1) Woodworth Wells number group checking test; (2) a simple addition and subtraction test; (3) directions test; (4) letter substitution; (5) spelling; (6) alphabetical filing. These tests were given and graded by the same method as that used in the employment office, and by the same examiners.

The results of the study are shown in the following chart. For the sake of simplicity, actual averages are omitted and group ranks given instead.

		Group Tested	No. in Group	Rank in Tests	Rank in Educ.	Rank in Age	Average Age
Applicants tested Employment Office	1.	*Recommended and hired*....	287	7	7	7	21.
	2.	Recommended and not hired.	137	3	3	8	21.1
	3.	Not recommended—not hired	173	8	9	9	21.9
	4.	Not recommended but hired.	19	9	8	6	20.8
Pupils tested in school	5.	High school girls...........	143	1	1.5	3	17.9
	6.	*High school boys*	38	5	1.5	2	17.1
	7.	X Business College (A. C.)...	31	2	6	1	16.2
	8.	X. Business College (B. C.)..	42	4	5	4.5	18.5
	9.	Y Business College.........	38	6	4	4.5	18.5

Rank in education is based on the number of years completed in school. The high school seniors had almost twelve years of schooling to their credit and are therefore ranked highest in this respect. The group of applicants not recommended and not hired had only nine years of schooling to its credit and is ranked last.

The results of this study showed very clearly:

1. That the senior girls of the high school stood highest in their performance. On the basis of actual averages the performance of this group in the tests was 30% better than the

performance of Group 1, the group of applicants who were tested in the employment office, recommended, and hired.

2. The group tested in the Employment Office and recommended but *not hired* (Group 2 on the chart) is the only group among those tested in the employment office which compares favorably with the high school girls and girls of the X business college (advanced class.)

3. The pupils of the Y business college and of the beginners class of the X business college compare favorably with Group 1, the group of applicants recommended and hired.

4. Though the number of high school boys tested is comparatively small, their performance was considerably poorer than that of their girl classmates.

5. There is the usual correlation between performance in tests and education, but a much less marked correspondence between rank in tests and age.

Some of the above facts raised pertinent questions as to the existing employment methods.

In the first place, if the girls recommended by the psychological section of the employment office but not hired, represented the highest group of applicants, why were they not employed? The interview records showed that thirty-five of these applicants were either teachers or high schools pupils who wanted only summer work. Other reasons for not employing these applicants were: (1) They were already employed elsewhere; (2) the wages they demanded were above the starting rate for clerical work; (3) positions were not available at the time. With regard to the last two reasons, it was obvious that better rates and better selection might be in order.

In the second place, why were not more candidates secured from the high school and business college graduating classes? If these girls were much better in the tests than those hired, it woud have paid to offer them higher wages than the existing standard maximum. Besides superiority in the tests, this group had the benefit of 2.3 years more of education. Moreover, they were 2.9 years younger than the girls hired, a fact which would lessen the likelihood of their leaving on account of marriage. The hiring of a superior type of clerk would also make it possible to reduce the amount of equipment and supervision required. Fewer desks, less space, less supervision, and fewer mistakes, all incident upon the employing of a better type of clerk, would warrant an increase in the wages paid.

In the third place, why were nineteen girls who were not recommended by the psychological section, hired nevertheless?

Nineteen out of 616 is, to be sure, a small number. Of this number four had left or been discharged as unsatisfactory. Why was the number of those leaving so small? Partly because a majority of these girls had been brought in at the suggestion of department heads or friends. It is difficult to check up the work of clerks hired in this way, largely because they are working among friends. However, it frequently happens that an unsatisfactory applicant is hired because of a pressing emergency.

To sum up, the results of this study showed the importance of some knowledge of the available labor supply, and a basis upon which to determine its relative desirability. It showed that predetermined starting wages had eliminated some of the most desirable applicants. It showed that a system of flexible starting wages was highly desirable, *if* the wages could be varied in accordance with the quantity and quality of the available labor supply. These problems are among the most important and recurrent which the employment manager has to deal with.

PICTURE COMPLETION

Robert H. Gault, Northwestern University

This paper reports an attempt at scoring picture completion records obtained from one hundred and eighteen delinquents and dependents in the Juvenile Psychopathic Institute in Chicago. The major portion of the work was done by Mrs. Ruth Sanders Thompson, Fellow in Northwestern University during the academic year 1918-1919, under the direction of the Department of Psychology.

To obtain the one hundred and eighteen cases, Mrs. Thompson read without omission, a block of nine hundred records that had been made prior to September 1918. These include all the records made at the Institute during a period of almost two years. From these were selected all those who had undergone test by the Binet and the Stanford Revision, the Seguin Form board, and Healy's Picture Completion, (older form) and whose intelligence quotient was sixty or over. There was no other selection. The chronological ages of the cases range from seven years and no months to fourteen years and eleven months. Only two of these are at age seven years to seven years and eleven months, and four at age eight years to eight years and eleven months. It happens that the remainder of the one hundred and eighteen cases are almost equally divided among the ages nine years to fourteen years and eleven months. The intelligence quotients of seventy-four of these one hundred eighteen Juveniles are almost uniformly distributed over the range of sixty-five to 79.9 inclusive. Sixteen lie between sixty and sixty-four and nine tenths; fourteen are between eighty and eighty-four and nine tenths; six between eighty-five and eighty-nine and nine tenths; five between ninety and ninety-four and nine tenths and the remaining three are between ninety-five and ninety-nine. The mode, therefore, lies between seventy and seventy-four and nine tenths; rather nearer the latter figure than the former.

The primary purpose in undertaking a study of these one hundred and eighteen was to retest them by such means as should be deemed suitable, and in the light of the data thus obtained and in the light, too, of their recent social history, to return to an examination of the original findings.

It is not my intention to enter into this aspect of the work in this paper. The secondary purpose of Mrs. Thompson's investigation was to find the degree of correlation among the records of the children in the tests referred to above viz: The Binet Tests, The Stanford Revision, The Form Board and the Picture Completion Test. It is with this aspect of the work, particularly as it relates to the Picture Completion test, that I am concerned here.

Naturally we first of all met the problem of scoring performance in the Completion Test. In the records of the Institute we found only an account of the moves made by each child: that blocks were placed here and there. We had before us Dr. Healy's method of scoring by "logical" and "illogical" errors, and Pinter and Anderson's method based upon the total frequency of choice of each of the blocks that were selected by their 1,438 presumably normal subjects. There is a possibility of 369 moves in this test. Our one hundred and eighteen cases made 360 of them as opposed to 349 made by Pintner & Anderson's normal subjects. A few of our moves, therefore, have received no score in their scheme. Some of these were of such a nature that one could conceive of situations in which these moves could merit Pintner and Anderson's minimum score but not more than that. Accordingly we have assigned arbitrarily, a score of 0.1 to these cases and graded the one hundred and eighteen by the Pintner and Anderson scale. Correlation with the Intelligence Quotient gives a result of 0.32.

In reviewing the Pintner and Anderson scores I am impressed with the thought that in a great many instances associations to which a very low score had been given to their subjects because of their infrequency, may really appear to the juvenile to be a reasonable solution of a practical problem. The criterion of reasonability among these younger folk is *already realized experience*: it may be oft repeated or infrequent, but striking.

Possible experience is, on their part, never, or at best very infrequently, a criterion of reasonability. It should not even require pointing out that the reasonability, rationality, or logical character of behavior is determined by the way in which one uses the data of experience at one's hand. One may take all obtainable data into account and nicely balance one's course in the light of all the details conflicting or otherwise. (One may put the broken window in the window space.) In that case we should place a high rating upon the behavior or adjustment and we should describe it as logical with distinction. We must remember that the phrase "all obtainable data" means one thing for me and quite a different thing for the youngster in the back streets.

One may overlook some of the details in the experience and suffer one's behavior or adjustment to the problem to be determined by the striking or unusual circumstance. In such cases, too, we should usually ascribe some degree of rationality and logical character to the behavior. If a youngster has seen a bottle of milk toppling from a high window but yesterday and narrowly missing a pedestrian below it would be a rational procedure, indeed I may say logical, for him to place the block bearing an *illustration of a milk bottle in the window* in the test and he should have liberal credit for making his own picture. If, on his way up and down the street one of the youngster's frequent visual experiences comes to him from the ubiquitous milk bottle enframed in each window in a row of tenements, once more it should be deemed rational and logical for him to place the bottle in the window space, and particularly so if he has, on rare occasions, taken a crack at it with his sling shot. That the connection-window space-milk bottle-occurred in but 0.3% of cases among Pintner and Anderson's subjects, and did not fall among the sixty-one most frequent cases, which they emphasize seems to me to be of less than great importance when one is attempting to build up a method of scoring that may be useful in estimating levels of intelligence.

I chanced to alight upon the milk bottle just now when I glanced around for illustrative material. The necktie-window, or the clock-window space connection or any one of a score of others, so varied is window furniture, embracing even socks, neckties and underdrawers, would have served my purpose quite as well.

Indeed I am sure that cases may occur in which the baby-fruit connection should be given a high grade for its logical character. Many a child and youth comes into our juvenile courts in large cities who knows nothing of picking fruit from a tree by the aid of a ladder and carefully placing it into a basket. Indeed we have not far to go to find youngsters even who do not know at all, or at best they know very vaguely, that fruit grows upon trees. Such a youngster may be dropping fruit upon the baby's head to tease him or may be bringing something pretty to comfort him. In either case, the reaction deserves a high grade for its logical and rational character. If you protest that any active and intelligent child will know and be able to appreciate the situation as you and I do, I can reply only with the counter protest that your judgment is a-priori and, in addition, assert my belief that thousands of city children have *at best* but a *vague* book knowledge of fruit-on-trees such as can hardly be said to have any value, even as second-hand experience, for directing behavior.

Let me say furthermore, that the infrequent youngster of inventive disposition, provided he approaches the task of completing the picture in a playful mood, as he should, will make unusual associations. It may appeal to his sense of humor to make the boy run away, not from the dog, but from the crying baby, and very well for who whether adult or juvenile, likes to be in the vicinity of a crying baby? Yet this connection would doubtless appear so seldom that, in Pintner and Anderson's scheme, no credit would be obtainable for it, or at best but a small credit, whereas, in the light of all the inner and outer circumstances, we should willingly give the performer a high grade for having made his own picture.

Of course, you have leaped to the conclusion, and rightly, that I have now in the back of my head, a picture of a certain opposition between the so-called "objective" and the so-called "subjective" attitudes toward psychology and a suspicion that in this particular instance, at any rate, the so-called "objective" attitude does not touch the psychological problem. Having brought our psychology from one direction to the study of the individual, we ought not immediately to depart from the individual in another direction. We do so, if at all, at the peril of sacrificing the significance of psychology in its relation to human behavior. If we pin ourselves to the mere frequency of a form of behavior and make that a criterion of high standing in a test we are bound to lose sight, in the shuffle, of that individuality that is so highly prized in the world because it makes things go.

And so, as I said above, we came to the need for a method of scoring our one hundred and eighteen performances in the picture completion test. Dr. Healy's "logical" and "illogical" or "rational" and "irrational" criteria gave us our starting point, and we were immediately before the question, "what is logical?" and "what is rational?"

We arbitrarily assigned a value of one hundred to the broken window in the window space and the same to each of the following associations: running boy-dog; wood cutter-log; fruit-basket; milk-sleeping cat; playing boys-football; flying bird-tree; girl-hat; and chicken-chicken. The remainder of the sixty-one situations were then turned over to each one of a group of eight including instructors and advanced students in the department of Psychology in Northwestern University, with the request that each one assign a value to each position. Each member of the group understood that he must needs keep in mind the possible motives of children such as come into the Juvenile Psychopathic Institute in a city like Chicago; this means that each one should keep before him the possible backgrounds of individual experience among such youngsters. We then brought our ratings to the table. In many instances the values assigned to a given position varied so little among members of the group that we considered it safe to strike a simple average to obtain what we would accept as a final value. In other instances where a considerable divergence obtained we discussed the case thoroughly at the conference table. There we sought for all possible light upon the ways in which a youngster, genius or of common mold, could interpret the situation in question. Finally we came to practical agreement upon the question of value in every case and the simple average gave us the accepted final weight. This done, it was a simple matter to apply our scale to the records of the one hundred and eighteen cases and to find the correlation of the standings so obtained with the intelligence quotients.

The correlation-Pearson coefficient-between the grades of the one hundred and eighteen in the picture completion and the I. Q. was 0.41. The result is considerably lower in relation to the Binet 1911 series.

We do not claim particular value for these results themselves. It is probable that we assigned too great weight to particular situations. Nor do we attach great importance to our method of obtaining a scale in as far as that method consists in sitting round a table in a suburban town seeking agreement of opinion as to what may have been the motives of back-street children in Chicago a year earlier or more when they placed a pot of flowers or a milk bottle in the window space in the picture. It does represent, however, an honest effort to take advantage of the psychologic background of the exceptional child.

I believe finally that the criteria employed by Dr. Healy are capable of development along the lines suggested in our crude experiment, and I have only this to propose that investigators, in their use of picture completion tests make every effort, by conversation with each individual under test to discover his motive for his several choices. When this has been done to the fullest extent possible, and when these motives have been studied in relation to the social history of each case, it will be possible ultimately to assign a weight to each choice or move varying with the types of experience represented. A given youth may receive a high grade for a certain performance and another a lower grade for the same reaction. What seems, to the adult placed as we are here, an illogical reaction may in the light of our research be logical.

I am aware that this suggestion points to a heavy tedious task but I am convinced that it offers one method of giving proper weight to experience in our psychological tests.

A COMPARATIVE STUDY OF THE INTELLIGENCE OF SEVENTY-FIVE WHITE AND FIFTY-FIVE COLORED COLLEGE STUDENTS BY THE STANFORD REVISION OF THE BINET-SIMON SCALE

By S. M. Derrick, M. A., University of South Carolina

The study of the comparative intelligence of the negro and white races has engaged the attention of students only in recent years. The first investigation of this kind was made at the University of South Carolina in 1913 by Miss Strong (14). Since then similar studies have been made which have in the main corroborated the results secured by her. The investigations made by Sunne (13), Pressey (9), Phillips (8), Perring (6), Pyle (10), Odom (5) and others show that age for age the negro student is younger mentally than the white.

These studies were made with the grammar grade pupils exclusively. The present study undertakes to extend the investigation to the college students of both races.[1] It shall be the purpose of this paper (1) to make a comparison of the ages of the two groups of college students, (2) to make a comparison of the intelligence, (3) to determine the correlation of mental ability with school success, (4) to determine the correlation of mental ability with environment, and (5) to determine the difference in the specific mental traits of the two races.

This study was made at the University of South Carolina. The seventy-five white college students tested were students of the University, the majority being members of the Freshman and Sophomore classes. They were unselected students with whom the examiner was not acquainted previous to the testing. In the main, the students were tested in their own rooms. This was found to be the most convenient place and was favorable to the correct attitude on the part of the subjects. Care was taken to exclude all other persons from the room and distractions were kept to a minimum. The testing was done quietly and only when a subject could be found who had the time to give freely. Usually from forty-five to sixty minutes were consumed, depending of course upon the reaction of the individual. The tests were entered upon with interest, if not with zest. All seemed at ease and enjoyed the time spent.

---

[1] The work of Strong has been of valuable assistance in this connection, as have been the works of the other students of the races, referred to above. Many valuable suggestions have been received from Terman (16) as to the technique of administering the test. The writer is also deeply grateful to Dr. Josiah Morse for his sympathetic interest and helpful criticisms in the making of this study.

The fifty-five colored students were secured at two colored institutions in Columbia; namely, Benedict College and Allen University. Subjects were secured from every class in college. However, due to the scarcity of college students it was found necessary to take a few men from the two upper classes of the high school. But these were all above the age of 16. Everything was done to make the work both pleasant and profitable. Favorable circumstances existed and the writer did not have any difficulty in getting the subjects to feel at ease. Marked interest was shown and enthusiastic endeavor was manifested. The work was thoroughly enjoyed, and to the authorities the writer is deeply grateful.

In this testing the Stanford Revision of the Binet-Simon Tests was used and the directions given by Terman were followed very closely. All responses were marked + or —. It was not possible to have a second person to keep a full record of all that the pupils said and did, but the examiner kept in his note-book as many as possible of the answers to the definitions, problem questions, difference between abstract words, and the interpretation of fables.

Age Comparison

No student below the age of 16 years was tested. Terman states that at this age the student has reached his maximum performance in the test, that is, after this stage of his development added years do not increase his intelligence quotient. The 75 white college students ranged from the age of 16 to 25 inclusive. However, these were unevenly distributed, the greater number lying at the lower end of the range. The median age for the entire number was 19.

The number and percentage of students at each age was as follows:

Age	16	17	18	19	20	21	22	23	24	25
Number	7	14	20	14	10	7	1	1	—	1
Per cent	9.3	18.7	26.7	18.7	13.3	9.3	1.3	1.3	0	1.3

It will be seen that 26.7% are 18 years of age; also, that there are just as many one year above the median age as there are one year below. It is interesting to note that only 4% were above the age of 21. Some of the students are retarded in their schooling, but it was practically impossible to ascertain how many, or to what extent. Some entered college from the third year of the high school, while some entered from the fourth. Not a few spent one year or more out of school engaged in work, due to the demands of war, while many had seen military service. These circumstances would necessarily make any attempt to determine the amount of scholastic retardation futile.

The distribution of the ages of the 55 colored students was through a wider range—from 16 to 35 inclusive. The number and percentage for each year follows:

Age	16	17	18	19	20	21	22	23	24	25
Number	6	0	1	3	9	4	6	3	2	1
Per cent	10.9	—	1.8	5.5	16.4	7.3	10.9	5.5	3.6	1.8

(Continued)

Age	26	27	28	29	30	31	32	33	34	35
Number	5	4	2	2	0	2	3	1	0	1
Per cent	9.1	7.3	3.6	3.6	—	3.6	5.5	1.8	—	1.8

The highest percentage of negroes found at any group was *16.4% at 20. The median for age is the same as the mode.*

As with the whites it was impossible to determine the amount of school retardation. In fact, the task would have been even more difficult. A few of the students tested were found to have been out of school for as long as ten years before attempting to enter college. Due to the economic and social conditions under which they lived, they were forced to work some and then go to school. So varied were the causes that kept them from college, that no standards could be devised by which the amount of retardation could be determined.

The median age of the whites is two years less than the median age of the negroes, and the average age of the whites is five years less than that of the negroes.

The oldest negro is ten years older than the oldest white; 34.5% of the negroes are above the age of the oldest white, while only 10.9% lie below the median age of the whites. Also, 73.4% of the whites are below the median age for the negroes. From these comparisons it is evident that the negro college student is older than the white.

Comparison of Intelligence Quotients

According to Terman (16) The average intelligence quotient of college students should be 115, the highest possible score being 128. The average for the 75 white college students tested was 112. Many of the students passed two or more of the "Superior Adult Test," and had there been higher tests they might have increased their intelligence quotients.

Of the total number of white students 40% tested above 112, the average, while 57% tested below. More students tested at 108 than at any other point, the number being 10, or 13% of the total. See table I, section A.

TABLE I

INTELLIGENCE QUOTIENTS OF WHITES AND NEGROES

(A) 75 White College Students (B) 55 Negro College Students

I. Q.	Number	Per Cent	I. Q.	Number	Per Cent
70- 80	0	0	70- 80	1	1.8
80- 90	0	0	80- 90	4	7.4
90-100	5	6.7	90-100	15	27.3
100-110	34	45.3	100-110	24	43.7
110-120	26	34.7	110-120	10	18.2
120-130	10	13.3	120-130	1	1.8

Average I. Q. = 112±6. Average I. Q. = 103±7.8

All the white college students tested at or above average intelligence. Of the 75 students 53% tested at average, and 47% above, 30.7% being between 105 and 110, 14.7% being between 100 and 105, and only one being below 100. Of the entire number 38.6% tested at superior intelligence and 14.7% at very superior.

When the intelligence quotients of the negro are analyzed the most striking thing is their wide distribution. They range from 76 to 125. The average for the negroes is 103 or 12 points below Terman's average intelligence quotient, and 9 points less than the average for the whites. The mean variation was 7.8.

Of the 55 tested, 9.2% tested below average intelligence. Thirty-nine or 71.0% tested between 90 and 110, average intelligence. However, there was a greater percentage in the upper half of this range, 23.6% having tested between 105 and 110, 16.4% between 100 and 105. Ten or 18.2% made a score of 108, the largest number at any one intelligence quotient. It is interesting to note further that 18.2% of the negroes tested at superior intelligence and 1.8% at very superior. A greater percentage of the negroes tested above average intelligence than below, 20.0% having tested above and 9.2% below.

More students from both races tested at 105-110 than at any other five-point group. Of all the whites 9.3% tested below the average for the negroes, 83.6% of the negroes tested below the average for the whites.

While the scale of intelligence quotients for the negroes extends lower than that of the whites, it ranges approximately as high. The highest score made by any colored student was 125; the highest made by any white student was 128. Odom

(5) looks upon this fact with optimism for the negro. He claims that the fact that the records of a limited number of negroes equal the records of the best whites gives indication of larger possibilities for the former.

It will be of interest to note the variability of the two races. Mayo (4) found that the negro was about 90% as variable as the white in scholastic attainment. In the present study the opposite is found to be the fact; the whites are only about 77% as variable as the negroes. This would indicate that we should expect to find more inferior people among the negroes than among the whites.

A comparison of the average intelligence quotients of the different age groups, and also of the mean variations is of interest at this point. The table below is a summary of the results obtained:

TABLE II

AVERAGE SCORE AND MEAN VARIATION OF EACH AGE GROUP

	(A) 75 White College Students			(B) 55 Negro College Students		
Age	Number	Average Score	Mean Variation	Number	Average Score	Mean Variation
16	7	110	6.6	6	114	4.7
17	14	110	5.1	0	0	0
18	20	113	5.4	1	94	0
19	14	111	7	3	109	2.3
20	10	111	4.8	9	102	5.1
21	7	113	5.6	4	100	10
22	1	108	0	6	95	8.3
23	1	100	0	3	99	9.3
24	0	0	0	2	104	0
25	1	111	0	1	107	0
26				5	102	5.4
27				4	109	3.5
28				2	110	2.5
29				2	105	2.5
30				0	0	0
31				2	106	5.5
32				3	94	8.3
33				1	94	0
34				0	0	0
35				1	108	0

No very pronounced difference exists between any age group. An 18 year old student made the lowest score of any of the whites, while the highest score was made by a student of 19. The mean variation of each age-group varies more than the average intelligence quotients of each group. The lowest is in the 20-year group where the average deviation is 4.8, the highest is in the 19-year group where it is 7.

A similar condition is noted among the negroes. However, the average score for each age-group of negroes varies through a wider range than does the average score for each age-group of whites. Also, the mean variation of the different groups vary widely, ranging from 2.3 to 10. See table II, section B.

One of the most outstanding facts is the comparative high score made by the 16-year old negroes. The average intelligence quotient of this group is two points higher than average for all the negroes, four points higher than the average for any other age-group of negroes one point higher than the average for any group of whites and two points higher than the average for all the whites. While the number is small the comparison is nevertheless interesting. The mean variation of the 16-year group of whites is higher than for the negroes, the figures being 6.6 and 4.7 respectively. The highest score made by any 16-year old white is 123, while the lowest is 101; the highest for the same age-group of negroes is 125, while the lowest is 107. It was a 16-year old student that made the highest score made by any negro student. The only other age group in which anything like a comparison can be made is the 20-year group. The average score of the negroes in this group is 102 with a mean variation of 5.7; the average score of the whites in this group is 111 with a mean variation of 4.8. However, when the average score of each age group of whites is compared with the average score of each similar age group of negroes it is found that in the main the whites rank higher than the negroes.

Correlation of School Success with Intelligence Quotients

Terman (15) classified the school success of 504 children into five grades: *very inferior, inferior, average, superior* and *very superior.* Similarly the college work of the 75 white students tested was classified. The grades for the making of this classification were secured from the Dean's Office of the University. The number and percentage in each group is here given in Table III.

TABLE III

Comparison of School and Test Rating (White)

Group	No. Students	School Rating	Test Rating	Average Score	Mean Variation
Very Inferior..........	6	8%	0	106	2
Inferior................	23	30.7	0	106	3.8
Average................	18	24.	52%	113	3.6
Superior..............	17	22.7	34.7	113	5.9
Very Superior.........	11	14.6	13.3	121	4.5

When this grouping is compared with that made on the basis of the intelligence quotients it is seen that the agreement below the average school success is not very close. In fact no white college student tested below average intelligence, while 29 or 38.7% are doing inferior and very inferior college work. However, a majority are doing average to very superior college work, the number being 46, or 61.3% of the total. It would seem that the college work of the students is below their capacity, as revealed by the test.

It will be seen that the average score of each of the three lower school success groups is too high, but that the superior and very superior show a very close correlation. One rather striking fact brought out was that the subject that made an intelligence quotient below 100 is doing superior college work. However, the one that scored highest in the test is doing very superior college work. There are some rather serious disagreements. Many of those that are doing inferior to very inferior college work tested relatively high. But it must be remembered that school success rests on other factors than intelligence, such as application, quality of instruction, health, motive and interest. In spite of these considerations the correlation between school success and the intelligence quotients is high, it being .70 according to Pearson's formula (12).

The school success of the negroes was secured from the authorities of the two negro institutions. The number and percentage in each group is given in the following table. In the third column are the percentages determined by the tests.

TABLE IV

COMPARISON OF SCHOOL AND TEST RATING (NEGRO)

Group	No. Students	School Rating	Test Rating	Average Score	Mean Variation
Very Inferior..........	3	5.4%	1.8	99	9.7
Inferior..............	9	16.4	7.4	99	6.6
Average..............	15	27.3	70.	99	7.5
Superior..............	18	32.7	18.2	107	5.3
Very Superior.........	10	18.2	1.8	109	5.1

The above figures indicate that the school work of the negroes is given a higher rating by the teachers than it deserves. An interesting fact is that the percentage that tested above average intelligence is nearly identical with the percentage classed as doing average college work.

No difference at all is found in the average score of the three lower groups, each being 99. However, it will be seen that the very inferior group is more variable than the other two groups. But the number is so small that the comparison could not at all be fair. Some rather serious disagreements are noted here. In only one group does the average score correspond to the classification based on school success, that is, the average group. But as a whole the correlation is found to be rather high, it being .50.

One important consideration must be borne in mind when a comparison is made between the two groups of students—the *difference between the standard of the negro institutions and* the standard of the white institution. According to the rating of the two negro institutions given in Bulletin 1916, No. 39, U. S. Bureau of Education, the standards required to be maintained by the negro students are inferior to those required to be maintained by the white students. On the other hand it is reasonable to suppose that the average negro who has the opportunity to attend college puts forth greater effort than the average white student, which may account for the fact that the negro's school work receives a rating superior to his intelligence quotient, whereas the white student's school work is inferior to his intelligence quotient.

In every school success group the average intelligence quotients of the whites is greater than that of the negroes, and the mean variation is less except in the superior group in which it is .6 higher. Further there is a lower degree of correlation between the school success and intelligence of the negroes than there is between the school success and intelligence of the whites.

Correlation of Environment with Intelligence Quotients

In making a comparison of the environments of the two races with the intelligence quotients it must be kept in mind that the average environment of the negro is much inferior to that of *the white.* The environment that would be considered average or even inferior for the white, would be considered superior for the negro. However, this fact will not prevent the making of some interesting comparisons.

The environments of the whites were classed into five groups; namely, *very inferior, inferior, average, superior* and *very superior.* Of the 75 tested, not one was found to have very inferior environment. The greatest percentage was found in the average group, there being 45.3%. Only 8% were classed as inferior; 36% were in the superior group and 10.7% in the very superior. Combining these figures we have 46.7%of the white college students from above average environment, with only 8% below.

When the intelligence quotients of the environmental groups are analyzed it is found that there is a rather high degree of correspondence. The coefficient of correlation is .85. Each higher environmental group has a higher average intelligence quotient than the preceding one. A summary of the results is given in table V.

TABLE V

COMPARISON OF ENVIRONMENTAL AND TEST RATING (WHITE)

Group	No. Students	Environ. Rating	Test Rating	Average Score	Mean Variation
Very Inferior.........	0				
Inferior..............	6	8%	0	103	5.7
Average..............	34	45.3	52.	111	5.9
Superior..............	27	36.	34.7	112	5.4
Very Superior........	8	10.7	13.3	118	4.0

The mean variation of each group does not vary widely from the variation of every other group. The smallest is in the very superior group, meaning that the students in this group test very superior to a greater degree than those in the superior environmental group test superior, those in the average group test average or those in the inferior group test inferior.

When the colored students were classified as to environment, none was found in the very inferior group, 7.3% were in the inferior group, 23.6% in the average, 52.7% in the superior and 17.3% in the very superior. It is a rather a striking fact that a comparatively small percentage of the negro students are classed as being from inferior or average environment, there being 70% from above the average negro environment. This seems to indicate that only those negroes from the best homes and most progressive communities go to college to any extent.

When the environment is correlated with the intelligence quotients the agreement is found to be rather close, the coefficient of correlation being .70. A summary of the results is given below in table VI.

TABLE VI

COMPARISON OF ENVIRONMENTAL AND TEST RATING (NEGRO)

Group	No. Students	Environ. Rating	Test Rating	Average Score	Mean Variation
Very Inferior.........	0		1.8		
Inferior..............	4	7.3%	7.4	98	8.0
Average..............	13	23.6	70.0	99	8.2
Superior..............	29	52.7	18.2	104	6.0
Very Superior........	9	17.3	1.8	110	6.6

The superior environmental group shows less variability than any other, while the average group shows the greatest. However, the range of the group variation is not very wide, it being from 6 to 8.2.

In every environmental group the average score of the white students is higher than that of the negroes and the mean variation is less. The negro college student is from a better negro environment than the white is from a white environment. Seventy per cent of the negroes are from above the average environment, while only 46% of the whites are from such. This indicates that white boys from the average white environment go to college nearly twice as frequently as do the negro boys from the average negro environment.

There is a higher degree of correlation between the environment and the intelligence quotients of the whites than there is between the environment and intelligence quotients of the negroes, the coefficients being .85 and .70 respectively.

The conclusions reached here are in accord with those of the other investigators in this field. Pyle (2) says that the good social group of the negroes stands about midway between the poor social group and the good social group of the whites. Ferguson (1) states that the negro boy from a good social group makes about three-fourths the showing of the white. Odom (5) and Phillips (8) in the main, corroborate this. Odom (5) say that " it would be clearly impossible for the negro to show the same manifestations of mental traits as whites after having been under the influence of an entirely different environment for many generations,"

Specific Differences

The responses of both races, both successes and failures, correspond to the specimen answers given by Terman (16) and others. The percentage of white and colored students that passed each specific test is given below for both races in table VII.

TABLE VII

	Negroes	Whites
YEAR XIV	55	75
Vocabulary, 50 words	87%	99%
Induction Test	78	96
Difference between President and King	96	92
Problems of Fact	95	99
Arithmetical Reasoning	89	99
Reversing Hand of Clock	75	85

TABLE VII—*Continued*

	Negroes	Whites
YEAR XVI	54	75
Vocabulary, 65 words	46%	60%
Interpretation of Fables	75	82
Difference between Abstract Words	89	74
Repeats 6 Digits Reversed	51	54
Code	22	60
YEAR XVIII	38	74
Vocabulary, 75 words	5%	21%
Binet's Paper Cutting Test	49	82
Repeats 8 Digits	29	32
Repeats Thought of Passages	10	44
Repeats 7 Digits Reversed	15	18
Ingenuity Test	25	63

Year XIV.

The negroes were more successful in giving the difference between a president and a king than in any other test, 96% giving at least two of the three differences correctly. Also the negroes did well with the problems of fact, 95% passing it correctly. Of all the tests in this group they did poorest in changing the hands of the clock. Sunne (13) also found this condition and suggests that it may be due to the lack of acquaintance with a timepiece. This, however, is hardly true for college negroes, who, as we have seen, come from superior negro environment. More probably it is due to the inability to reason the problem out.

Practically all the white students were able to pass all the tests of this group. Reversing the hands of the clock was found to be the most difficult, 85% being able to pass it. Over 90% were able to pass the other tests.

Year XVI

This group was not given to one of the negroes, due to the fact that he failed to pass more than one of the tests in the preceding year group. The negroes were most successful in giving the difference between abstract words, 89% having been able to pass this test. They were least successful with the code, only 22% doing it well enough to make a pass. They were also rather successful in the interpretation of fables, 75% interpreting four of the five correctly.

The interpretation of fables proved to be the easiest for the whites in this group, 82% passing it; while repeating six digits reversed proved to be the most difficult with 54% repeating them correctly.

Year XVIII

The greatest success was with Binet's paper cutting test, 49% of the negroes doing it correctly. It is a rather striking fact that only 5% were able to pass the vocabulary test of this group.

Binet's paper cutting test proved easiest also for the whites, 82% passing it. The repeating of seven digits reversed was the most difficult, only 18% repeating them correctly.

In every test the percentage of passes registered by the whites is greater than the percentage registered by the negroes, except in giving the difference between a president and a king, where the negroes excelled the whites by 4%; and giving the difference between abstract words, where the whites are excelled by 15%. This would seem to indicate that the negro is superior to the white in discrimination and abstraction. However, it seems reasonable to suppose that in the giving of the difference between abstract words the peculiar training of the negro has determined his success. It has been drilled into him by his teachers that there is a decided difference between *poverty* and *misery, laziness* and *idleness,* and *reputation* and *character.* No doubt that training also enters into his success in the interpretation of fables. It is characteristic for the negro to teach and preach by the use of moralizing stories and fables.

The conclusion of Pressey and Teter (9) that the negro is *better in concrete and routine problems than in problems* involving mental abstractions or reconstruction is found to be generally true in this study. The conclusion of Sunne (13) that colored children find it difficult to reverse the hands of the clock is corroborated, but his conclusion that the problems of fact were difficult is not. The code was not found to be easier for the negroes than for the whites as is stated by the same investigator. The conclusion of Phillips (8) that the negro is slower in response and less animated was not corroborated by this study. Sunne (13) could find no such difference.

One of the most striking things brought out in the testing of the negroes was the answers given to the first problem in the "Problems of Fact" test. Fifty per cent of them either said that a man was lynched or a man was shot by a mob. Very few such answers were given by the white students. This fact would tell an interesting story to a psycho-analyst.

Summary

1. The negro college student is older than the white; and according to chronological age he is more retarded.

2. The white college students have a higher average intelligence quotient than the negroes, their average score being 9 points higher.

3. The range of the intelligence quotients of the negroes was wider than that of the whites, the range of the former was from 76 to 125 and that of the latter was from 91 to 128.

4. The white college student was about 77% as variable as the negro.

5. The average intelligence quotient of the 16 year old negroes is higher than average for any other age group of either race.

6. The college work of the white college student is inferior to his ability as measured by the test.

7. The college work of the negroes is given a higher rating by the teachers than it deserves according to the test.

8. In every school success group the average intelligence quotient of the whites is higher than that of the negroes, and the mean variation is less in every group except in the superior group where it is .6 higher.

9. There is lower degree of correlation between the school success and intelligence quotients of the negroes than between the school success and intelligence quotients of the whites; the coefficient of correlation of the former being .50, that of the latter being .70.

10. In every environmental group the average score of the whites is higher than that of the negroes, and the mean variation is less.

11. White boys from average environment go to college nearly twice as frequently as negro boys from average negro environment.

12. There is a higher degree of correlation between the environment and intelligence quotients of the whites than there is between the environment and intelligence quotients of the negroes, the coefficients of correlation being .85 and .70 respectively.

13. From this study it seems that the negro is better in memory and in concrete and routine problems than in those that involve mental abstraction and reconstruction.

Bibliography

1. Ferguson, G. C. Jr. The Psychology of the Negro. *Archives of Psychology*, No. 36, April, 1916.
2. Gatewood, E. L. Teacher's Estimates of Negroes and Whites. *School and Society*, Vol. IX, 1919, pp. 89-90.
3. Redfern, L. H. and Rich, S. G. Binet Tests on South African Natives—Zulus. *Pedagogical Seminary*, Vol. XXIV, 1917, No. 3, pp. 373-383.
4. Mayo, W. J. The Mental Capacity of the American Negro. *Archives of Psychology*, No. 28, November, 1913.
5. Odom, H. W. Negro Children in the Public Schools of Philadelphia. *American Academy of Political and Social Science*, Vol. XLIX, pp. 186-208.
6. Perring, L. F. A Study of the Comparative Retardation of Negro and White Pupils in a Philadelphia School. *Psychological Clinic*, Vol. IX, No. 3, pp. 87-93.
7. Phillips, B. A. Retardation in the Elementary Schools of Philadelphia. *Psychological Clinic*, Vol. VI, No. 3, pp. 79-81.
8. The Binet Tests Applied to Colored Children. *Psychological Clinic*, Vol. VIII, No. 7, pp.190-197.
9. Pressey, S. L. and Teter, G. F. A Comparison of Colored and White Children by Means of a Group Scale of Intelligence. *Journal of Applied Psychology*, Vol. III, 1919. No. 3, pp. 279-282.
10. Pyle, W. H. The Learning Capacity of the Negro. *Psychological Bulletin*, Vol. XIII, 1916, pp. 82-83.
11. Pyle, W. H. The Mentality of the Negro Compared with the Whites. *Psychological Bulletin*, Vol. XII, p. 71.
12. Starch, Daniel. *Experiments in Educational Psychology.* The Macmillan Company, New York, 1919.
13. Sunne, Dagny. A comparative Study of Negro and White Children. *Journal of Applied Psychology*, Vol. I, 1917, No. 1, pp. 71-84.
14. Strong, A. C. Three Hundred Fifty White and Colored Children Measured by the Binet-Simon Scale of Intelligence: A Comparative Study. *Pedagogical Seminary*, Vol. XX, 1913, pp. 485-515.
15. Terman, L. M. and others. Stanford Revision Binet-Simon Scale and Some Results from its Application to 1,000 Non-Selected Children. *Journal of Educational Psychology*, Vol. VI. 1915. No. 9, pp. 551-562.
16. Terman, L. M. The Measurement of Intelligence. Houghton Mifflin Company, New York, 1916.
17. Some Suggestions Relative to the Study of the Mental Attitude of the Negro. *Pedagogical Seminary*, Vol. XXI, 1914, No. 1, pp. 199-203.
18. Woodworth, R. S. Comparative Psychology of the Races. *Psychological Bulletin*, Vol. XIII, 1916, pp. 388-396.

THE EFFECT OF CLIMAX AND ANTICLIMAX ORDER OF PRESENTATION ON MEMORY

By HENRY FOSTER ADAMS, University of Michigan

In previous papers,[1] the memory efficiency of size of advertisements, of their frequency of presentation with both duplication and variation, and of mixed sizes of advertisements, have been discussed. In the last paper mentioned, it was found that the memory value of the same total amount of space varied as it was made up of different combinations of sizes. For instance a series made up of 3 half and 1 quarter page advertisements of the same commodity was found to have a greater memorability than a similar series composed of 1 full and 3 quarter pages, yet the total area is the same in both cases, 1.75 pages, and the number of presentations is 4 in each case.

But a question which was left unsolved in that experiment was whether the order in which the mixed sizes were presented to the subject had any effect upon the memorability of the series. For example, if a firm advertises 4 times with mixed sizes of advertisements, is it more effective upon memory to present the larger sized advertisements at the beginning or end of the series.

To test the point 3 dummies were prepared, as follows. In dummy A, for instance, each firm advertised 4 times, the different insertions being variations and at least two sizes of advertisements being included in the 4 insertions. Each different arrangement of sizes was likewise represented by four different firms. With two of the firms, the larger sizes appeared first, followed by the smaller ones and with the other two firms for the same combination of sizes, the advertisements of smaller area appeared first, followed by the larger ones. In all, the following combinations were presented:—

---

[1] The Relative Importance of Size and Frequency in Forming Associations, *J. of Phil., Psy., and Sci. Meth.*, Vol. XII, 1915, pp. 477-491.

The Relative Memory and Values of Duplication and Variation in Advertising, *ibid.*, Vol. XIII, 1916, pp. 141-152.

The Memory of Mixed Sizes of Advertisements, *J. of Exper. Psychol.* Vol. II, 1917, pp. 448-465.

Dummy	A	B	C	
FFFH	2	2	2	
HFFF	2	2	2	
FFFQ	2	2	2	
QFFF	2	2	2	
FFHH	2	2	2	
HHFF	2	2	2	
FFQQ	2	2	2	
QQFF	2	2	2	
FFHQ	2	2	2	
QHFF	2	2	2	
FHHH	2	2	2	
HHHF	2	2	2	
FQQQ	2	2	2	
QQQF	2	2	2	
FHHQ	2	2	2	
QHHF	2	2	2	
FHQQ	2	2	2	
QQHF	2	2	2	
HHHQ	2	2	2	
QHHH	2	2	2	
HHQQ	2	2	2	
QQHH	2	2	2	
HQQQ	2	2	2	
QQQH	2	2	2	
	48	48	48	144

In the experiment, part of the subjects began reading the dummy at page 1 and went along progressively to the end; others began at the end and read back to page 1. This was easily accomplished by turning the dummy upside down. Therefore, to those subjects who read the dummy through in the forward direction, the advertisements of a certain firm would be seen in the order FFFH; while to those who read it through in the reverse direction, the same advertisements of the same firm would appear in the order HFFF. In this way each series of advertisements was seen by the subjects in both the climax and anti-climax order. For the sake of convenience of expression, the series of advertisements is said to be in the climax order if the smaller spaces appear first followed by the larger ones; in the anti-climax order if the larger spaces are seen first followed by the smaller ones.

With dummy A, 93 subjects were used in the direct or forward order of reading, 68 in the reverse order; in dummy B 95 in the direct order, 58 in the reverse; in dummy C 79 in the forward order and 70 in the reverse order. This makes a total of 463 subjects.

Since the results of the experiments may be considered in several ways and the conclusions to be drawn from the experiment depend in part upon the method of treatment, the original data will be presented. It is given in Table I.

TABLE I

Order of Ads.	Dummy A				Dummy B				Dummy C			
	D	R	Ant	Cli	D	R	Ant	Cli	D	R	Ant	Cli
FFFH........	137	103	137	103	53	29	53	29	92	76	92	76
	78	64	78	64	48	21	48	21	48	52	48	52
HFFF........	70	65	65	70	49	30	30	49	72	71	71	72
	70	53	53	70	58	34	34	58	69	92	92	69
			333	307			165	157			303	269
FFFQ.	74	56	74	56	49	26	49	26	76	81	76	81
	121	82	121	82	38	18	38	18	84	71	84	71
QFFF.........	104	86	86	104	36	11	11	36	32	34	34	32
	51	37	37	51	53	28	28	53	52	49	49	52
			318	293			126	133			243	236
FFHH.........	55	42	55	42	44	22	44	22	66	81	66	81
	71	62	71	62	50	22	50	22	17	35	17	35
HHFF..........	91	66	66	91	83	50	50	83	22	37	37	22
	110	91	91	110	23	11	11	23	87	89	89	87
			283	305			155	150			209	225
FFQQ.........	23	22	23	22	57	15	57	15	87	101	87	101
	115	87	115	87	57	29	57	29	95	81	95	81
QQFF.........	28	39	39	28	33	22	22	33	107	113	113	107
	32	38	38	32	100	61	61	100	35	47	47	35
			215	169			197	177			342	324
FFHQ.........	77	51	77	51	27	5	27	5	22	22	22	22
	57	45	57	45	114	54	114	54	36	57	36	57
QHFF.........	76	61	61	76	30	15	15	30	89	95	95	89
	36	30	30	36	37	17	17	37	55	67	67	55
			225	208			173	126			220	223
FHHH.........	60	41	60	41	59	26	59	26	18	27	18	27
	65	36	65	36	101	61	101	61	118	110	118	110
HHHF.........	88	57	57	88	18	7	7	18	106	91	91	106
	60	63	63	60	55	27	27	55	61	53	53	61
			245	225			194	160			280	304
FQQQ.........	70	68	70	68	65	30	65	30	25	23	25	23
	24	10	24	10	45	10	45	19	59	61	59	61
QQQF.........	81	28	28	31	74	43	43	74	44	65	65	44
	50	36	36	50	52	21	21	52	96	87	87	96
			155	159			174	176			236	224
FHHQ.........	27	26	27	26	24	10	24	10	59	62	59	62
	33	26	33	26	105	43	105	43	41	50	41	50
QHHF.........	50	49	49	50	63	34	34	63	104	83	83	104
	9	12	12	9	13	8	8	13	35	46	46	35
			121	111			171	129			229	251
FHQQ.........	42	38	42	38	37	21	37	21	77	78	77	78
	60	40	60	40	97	48	97	48	92	88	92	88
QQHF.........	91	84	84	91	29	17	17	29	52	70	70	52
	35	40	40	35	66	28	28	66	12	15	15	12
			226	204			170	164			254	230

TABLE I—*Continued*

Order of Ads.	Dummy A D	Dummy A R	Dummy A Ant	Dummy A Cli	Dummy B D	Dummy B R	Dummy B Ant	Dummy B Cli	Dummy C D	Dummy C R	Dummy C Ant	Dummy C Cli
HHHQ........	72	54	72	54	16	4	16	4	64	50	64	50
	58	40	58	40	20	14	20	14	66	62	66	62
QHHH........	21	22	22	21	79	30	30	79	62	65	65	62
	24	15	15	24	33	19	19	33	31	51	51	31
			167	139			85	130			246	205
HHQQ.........	5	10	5	10	8	6	8	6	39	52	39	52
	71	46	71	46	71	23	71	23	30	45	30	45
QQHH.........	11	10	10	11	48	16	16	48	86	79	79	86
	11	10	10	11	16	8	8	16	36	28	28	36
			106	78			103	93			166	219
HQQQ.........	9	12	9	12	18	5	18	5	18	22	18	22
	17	8	17	8	87	47	87	47	15	25	15	25
QQQH.........	12	23	23	12	18	9	9	18	65	37	37	65
	5	10	10	5	83	46	46	83	17	22	22	17
			59	37			160	153			92	129

The column of initials at the left gives the order of presentation of the advertisements, F standing for full page, H for half page and Q for quarter page. The second, labeled D, gives the gross credits received by the various series of advertisements when the subjects read in the forward or direct order, starting at page 1. The third column, labeled R gives the gross credits received by the advertisements when the dummy was read through in the reverse direction. The fourth column, headed by Ant. gives the credits for the advertisements which were read through in the anti-climax order; and in the fifth column, headed Cli. the credits for those advertisements which were read through in the climax order. The other columns gives the same data for dummies B and C.

TABLE II

	Dummy A Anti	Dummy A Cli	Dummy B Anti	Dummy B Cli	Dummy C Anti	Dummy C Cli	Total Anti	Total Cli
FFFH......................	333	307	165	157	303	269	801	733
FFFQ......................	318	293	126	133	243	236	687	662
FFHH......................	283	305	155	150	209	225	647	680
FFQQ......................	215	169	197	177	342	324	754	670
FFHQ......................	225	208	173	126	220	223	618	557
FHHH......................	245	225	194	160	280	304	719	689
FQQQ......................	155	159	174	175	236	224	565	558
FHHQ......................	121	111	171	129	229	251	521	491
FHQQ......................	226	204	179	164	254	230	659	598
HHHQ......................	167	139	85	130	246	205	498	474
HHQQ......................	106	78	103	93	166	219	375	390
HQQQ......................	59	37	160	153	92	129	311	319
	2453	2235	1882	1747	2820	2839	7155 105	6821 : 100

In Table II, a summary of the climax and anti-climax effects is given, indicating that the anti-climax order of presentation is somewhat more effective as far as memory is concerned.

But it must be remembered that these are gross figures and furthermore that unequal numbers of subjects were used in the forward and backward turning of the pages of the dummy. To compensate for this the credits for each advertisement were put into percentages which were multiplied

TABLE III

Order of Ads.	A				B				C			
	D	R	Ant	Cli	D	R	Ant	Cli	D	R	Ant	Cli
FFFH..........	529	494	529	494	218	244	218	244	332	262	332	262
	302	306	302	306	198	176	198	176	174	179	174	179
HFFF..........	270	312	312	270	201	252	252	201	259	245	245	259
	270	254	254	270	238	286	286	238	248	318	318	248
			1397	1340			954	859			1069	948
FFFQ..........	286	268	286	268	202	218	202	218	275	280	275	280
	466	393	466	393	156	151	156	151	304	245	304	245
	402	412	412	402	148	92	92	148	108	118	118	108
	198	177	177	198	218	235	235	218	187	170	170	187
			1341	1261			685	735			867	820
FFHH..........	213	201	213	201	181	185	181	185	238	280	238	280
	275	296	275	296	206	185	206	185	61	121	61	121
HHFF..........	352	316	316	352	341	420	420	341	79	128	128	79
	425	435	435	425	94	92	92	94	314	308	308	314
			1239	1274			899	805			735	794
FFQQ..........	89	102	89	102	235	126	235	126	314	350	314	350
	445	415	445	415	235	244	235	244	344	280	344	280
QQFF..........	104	187	187	104	135	185	185	135	385	390	390	385
	124	182	182	124	410	512	512	410	126	163	163	126
			903	745			1167	915			1211	1141
FFHQ..........	298	244	298	244	106	42	106	42	79	76	79	76
	221	215	221	215	469	454	469	454	130	196	130	196
QHFF..........	293	293	293	293	123	126	126	123	320	338	338	320
	149	144	144	149	152	143	143	152	199	232	232	199
			956	901			844	771			779	791
FHHH..........	232	198	232	198	243	218	243	218	65	93	65	93
	251	172	251	172	420	511	420	511	426	380	426	380
HHHF..........	340	264	264	340	74	59	59	74	381	315	315	381
	232	302	302	232	226	227	227	226	219	183	183	219
			1049	942			949	1029			989	1073
FQQQ..........	270	325	270	325	268	251	268	251	90	80	90	80
	93	48	93	48	185	159	185	159	214	210	214	210
QQQF..........	119	134	134	119	304	362	362	304	158	224	224	158
	193	172	172	193	214	177	177	214	345	301	301	345
			669	685			992	928			829	793
FHHQ..........	102	124	102	124	99	84	99	84	214	213	214	213
	128	124	128	124	431	362	431	362	148	172	148	172
QHHF..........	193	234	234	193	259	285	285	259	374	287	287	374
	35	58	58	35	53	67	67	53	126	159	159	126
			522	476			882	758			808	885

TABLE III—*Continued*

Order of Ads.	A				B				C			
	D	R	Ant	Cli	D	R	Ant	Cli	D	R	Ant	Cli
FHQQ..........	162	182	162	182	152	176	152	176	278	269	278	269
	232	191	232	191	398	403	398	403	332	303	332	303
QQHF..........	351	403	403	351	119	143	143	119	187	242	242	187
	135	191	191	135	271	235	235	271	43	52	52	43
			988	859			928	969			904	802
HHHQ..........	279	258	279	258	66	33	66	33	231	172	231	172
	224	191	224	191	82	117	82	117	239	213	239	213
QHHH..........	81	103	103	81	324	252	252	324	223	225	225	223
	93	72	72	93	135	160	160	135	111	176	176	111
			678	623			560	609			871	719
HHQQ..........	20	48	20	48	33	50.	33	50	141	179	141	179
	274	220	274	220	292	193	292	193	108	155	108	155
QQHH..........	42	48	48	42	197	135	135	197	300	273	273	301
	42	48	48	42	66	67	67	66	129	97	97	129
			390	352			527	506			619	764
HQQQ..........	35	57	35	57	74	42	74	42	65	76	65	76
	66	38	66	38	358	394	358	394	54	86	54	86
QQQH..........	46	110	110	46	74	76	76	74	234	128	128	234
	19	48	48	19	340	387	387	340	61	76	76	61
			259	160			895	850			323	457

by 100 to avoid the use of decimal points. This data is presented complete in Table III, and in summarized form in Table IV. Again we have a suggestion that the anti-climax order is somewhat more efficient than in the climax.

TABLE IV

	Dummy A		Dummy B		Dummy C		Total	
	Anti	Cli	Anti	Cli	Anti	Cli	Anti	Cli
FFFH....................	1397	1340	954	859	1069	948	3420	3147
FFFQ....................	1341	1261	685	735	867	820	2893	2816
FFHH....................	1239	1274	899	805	735	794	2873	2873
FFQQ....................	903	745	1167	915	1211	1141	3281	2801
FFHQ....................	956	901	844	771	779	791	2579	2463
FHHH....................	1049	942	949	1029	989	1073	2987	3044
FQQQ....................	669	685	992	928	829	793	2490	2406
FHHQ....................	522	476	882	758	808	885	2212	2119
FHQQ....................	988	859	928	969	904	802	2820	2630
HHHQ....................	678	623	560	609	871	719	2109	1951
HHQQ....................	390	352	527	506	619	764	1536	1622
HQQQ....................	259	160	895	850	323	457	1477	1467
							30677 105 : 100	29339

But again a difficulty of interpretation creeps in and this time it comes from the sheer magnitude of the figures themselves. The difference between 66 and 33 for instance is 33 or 100%; the difference between 540 and 507 is again 33 but only about 6%. There is to be found in the figures a dif-

TABLE V

	A	B	C	Total	Sum	Av.
FFFH	1071	895	1268	3234		
	988	1131	973	3092		
	1155	1253	947	3355		
	942	1201	1281	3424		
					13105	3276
FFFQ	1068	928	982	2978		
	1188	1032	1241	3461		
	1025	621	1092	2738		
	895	1079	910	2884		
					12061	3015
FFHH	1060	978	850	2888		
	930	1113	504	2547		
	899	1232	1621	3752		
	1022	980	983	2985		
					12172	3043
FFQQ	874	1866	897	3637		
	1072	965	1228	3265		
	1800	1370	1012	4182		
	1469	1249	1292	4010		
					15094	3774
FFHQ	1221	2528	1040	4789		
	1028	1034	664	2726		
	1000	1023	1058	3081		
	967	941	1165	3073		
					13669	3417
FHHH	1171	1114	700	2985		
	1460	822	1121	3403		
	778	798	827	2403		
	1301	1004	837	3142		
					11933	2983
FQQQ	832	1069	1125	3026		
	1939	1164	1020	4123		
	1127	1191	1419	3737		
	892	827	873	2592		
					13478	3370
FHHQ	823	1180	1005	8008		
	1032	1190	861	3083		
	1213	1100	767	3080		
	1655	1265	1261	4181		
					13352	3338
FHQQ	889	865	1031	2785		
	1214	987	1095	3296		
	1150	1200	1292	3642		
	1413	866	1210	3489		
					13212	3303
HHHQ	1080	2000	1342	4422		
	1172	700	1122	2994		
	1274	778	1010	3062		
	3574	1186	1586	3546		
					14024	3506

TABLE V—*Continued*

	A	B	C	Total	Sum	Av.
HHQQ..........	417	660	788	1865		
	1245	1510	697	3452		
	1142	685	907	2734		
	1142	1015	750	2907		
					10958	2739
HQQQ...........	614	1760	855	3229		
	1736	910	629	3275		
	2390	1028	548	3966		
	2522	1138	1246	4906		
					15376	3844
Sums and Av. of Averages					1100	3301
					1100	1100

ference between the absolute amount and the relative amount of influence resulting from the changing order of presentation.

Table of Ratios. The climax order is given an arbitrary value of 1,000 and the anti-climax a relative value. Since the climax order has always the value 1,000, it is omitted from the table. Table V was prepared, in which the anti-climax values are given in ratios, the climax order of presentation

TABLE VI

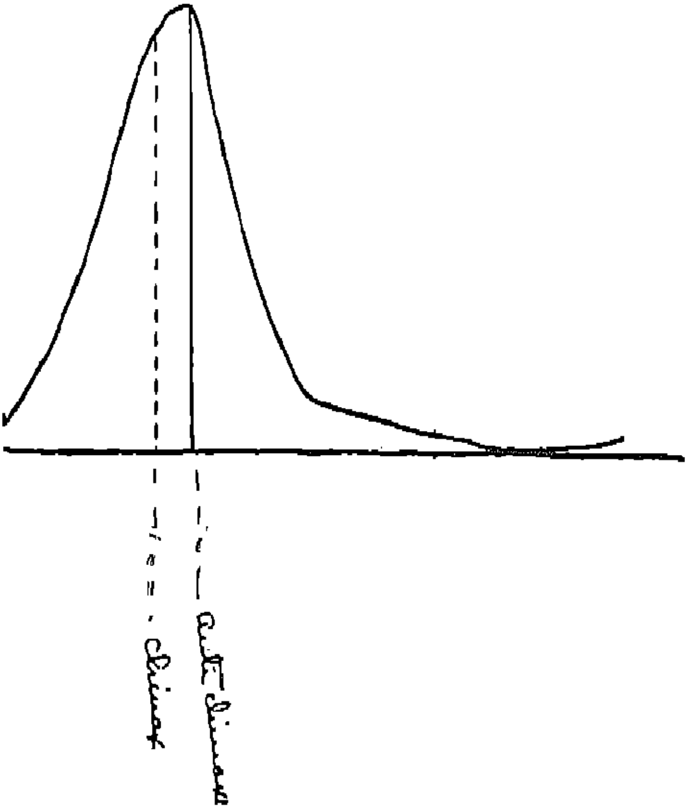
Anti-climax ratios put in the form of a frequency table.

V	F
40–59	3
60–79	15
80–99	37
100–119	49
120–139	23
140–159	6
160–179	4
180–199	3
200–219	1
220–239	1
240–259	2

Av. = 111.11
A. D. = 22.67
n = 144
P. E. = 1.6–

having an arbitrary value of 1,000. This table in summary indicates that the anti-climax order is 10% more effective than the climax order, and this fact is emphasized by throwing the anti-climax values into a table of frequency as is done in Table VI. This method of presenting the data gives the anti-climax order a value of 11+% advantage over the climax. The probable error is very low, only 1.6, approximately 1-7 of the significant difference. It may be concluded then, that the anti-climax order is more effective than the climax order.

GRAPH DRAWN FROM TABLE VI



THE SELECTION OF MILL WORKERS BY MENTAL TESTS.

By ARTHUR S. OTIS, *Camp Grant, Ill.*

The extended use of intelligence testing in the United States Army has given psychology—mental testing in particular—a widespread interest. It has come to the attention of many progressive industrial managers that a number of psychologists have been engaged for a number of years in research in the making and perfecting of tests for industrial employees, especially salesmen and clerks. The success of their work is attested to by the continued and extended use of their tests. The natural inference of both psychologists and industrial managers has been that ultimately psychological tests would be used to assist in choosing factory and mill workers.

The writer was called at the close of his army service (Jan. 1919) to the office of a large and progressive silk manufacturing company in Connecticut to "install a system of mental tests for prospective employees," (both clerks and mill workers). Bringing to bear his experience in mental testing in the army and schools the writer constructed two systems of intelligence tests, one for clerks and executives, similar to the Otis General Intelligence Examination, and a comprehensive performance intelligence examination somewhat similar to the performance tests used in the army for foreigners and illiterates. This latter was for all mill workers, a large percentage of whom were foreign or illiterate.

The clerical intelligence test proved to be of decided value in the selection of members of the office force and has been adopted for permanent use. It is not the purpose of this article to discuss this test.

The results of an extended tryout of the performance intelligence examination, however, are both startling and baffling and cast an ominous shadow over the future of strictly intelligence tests for manual workers.

The performance scale used in this connection consisted of 13 tests covering a wide range of mental and manual activity. The examination of an individual lasted on the average an hour. In all some 400 employees were tested. These were placed in three groups according to their productive ability ascertained by careful investigation. In the last analysis it was found that the correlation between intelligence and productive ability was zero! No amount of age grouping or length of service grouping would bring anything but zero out of the correlations.

That the tests did measure intelligence one may be perfectly confident, first because for the most part they were standard tests which have been used in one form or another for years—picture puzzle, drawing designs, paper folding, etc. The ability of clerks in these tests were found to be distinctly above that of mill workers as a class. Secondly the intercorrelations between the several tests ranged between .40 and .75 denoting a "reliability coefficient" for the whole scale of .97. When 13 widely varying tests tend strongly to measure the same ability that ability must be "general ability" or intelligence.

The writer is not prepared to say that it is proven conclusively that there is no correlation between intelligence and efficiency in mill work, for various reasons. It cannot be asserted confidently that every examinee did his best on the tests. There was lacking the incentive that is present when an applicant seeks employment. These examinees were already employed and saw no purpose in the tests. Moreover, the rating in productive ability was not as refined as could be desired due to numerous difficulties which could not be overcome in making the ratings. Furthermore, it was not possible to test such individuals as had failed to learn to weave or spin or be worth employing in some branch of the industry. These might have shown a lesser intelligence as a group.

On the other hand the difficulties above mentioned are only such as would be expected to "attenuate" the correlations—to reduce the true correlations to, say, one half or less. But a correlation must be pretty small in the first place if one half or one third of it equals zero! We are forced to conclude therefore that intelligence is a very unimportant factor in efficiency in silk mill work.

The psychologist can hardly rest content with such a conclusion without some sort of explanation, considering the known correlation between intelligence and clerical ability or executive ability or scholastic ability, military ability, etc. The following is suggested as a possible explanation.

Firstly, the machinery of the present day is so efficient and automatic that very little intelligence is needed to use it. Moreover, machine work is so specialized and monotonous that an intelligent person is very likely to revolt at it after a moderate period. The writer met one person who said she had sat in the same chair every work day for 18 years at her occupation of pasting strips of white paper around the edges of cardboard boxes. Yet she seemed entirely contented with the work and said she liked it. Some weavers had worked in the same mill for 35 years. Many persons do nothing but put full spools of silk in place of empty ones, week after week, year after year, with seeming contentment. Others do nothing but cut imperfections from thread and tie the ends—the imperfections being located by electric machinery automatically. Others do nothing but tie knots in breaking thread. Weaving is perhaps the most skilled operation in the mill. The operation of a loom is probably no more complicated than the operation of an automobile, yet three months are devoted to teaching a "weaver learner." It would seem that almost any degree of intelligence could be taught in that time to thread bobbins, tie knots, watch for imperfections, etc. The man who was considered the best weaver in the mill, having never failed to "make bonus" for a long period, took something like ten minutes to assemble the parts of a simple picture puzzle of twelve square pieces which intelligent persons had assembled in 25 seconds, and he was then satisfied to leave the ear under the chin. Another good weaver could not put the mouth under the nose in 25 minutes!

The conclusion which the writer draws from these researches is that intelligence is not only not required in a modern silk mill for most operations but may even be a detriment to steady efficient routine work. What qualities are required remain to be sought. Whether they are measurable is doubtful. They may be stolidity, patience, inertia of attention, regularity of habits, etc.

The field of industrial psychology as applied to manual labor is believed by the writer to be virgin soil.

THE COMMON SYNAESTHESIA OF MUSIC

By E. LEIGH MUDGE, State Normal School, Edinboro, Pa.

Synaesthesia, the reaction of one sense to the adequate stimulus of another sense, is not so uncommon as is often supposed. It is true that the very noticeable types commonly described are relatively rare, though even these are found by a recent writer to appear in from 9 to 15 per cent. of adults and in 25 per cent. of adolescents.[1] Like most tendencies in nature, synaesthesia appears in many degrees of variation. It is not a clearly-cut type of abnormality. It is, in fact, too common to be called an abnormality at all. Like any tendency, it is found in abnormal and extreme cases, but it has no absolutely definable lower limit. It is theoretically doubtful if any sense reaction takes place in absolute isolation from other senses. In the case of any meaningful experience of vision or audition, sight or hearing is the key that unlocks a complex of sensations or images or both in terms of a variety of senses, especially the deep-seated organic and strain senses and others intimately involved in feeling and emotive processes.

Doubtless there is a closer connection between either vision or audition and the deep-seated intimate senses than between these two articulative senses themselves. However, it is not strange if there be some tendency, with individual variations, to connect visual imagery with auditory stimuli. Various unusual cases of "colored hearing" and other inter-associations of vision with hearing have been reported. Reports from fifty mature students, chiefly active public school teachers, seem to indicate that the unusual cases represent an uncommon development of a common tendency. These fifty students were asked to report the colors or brightnesses which they associate with certain tones, keys, instruments, and familiar musical compositions. Only eight reported a total lack of such associations. While there is little uniformity as to particular color associations, some tentative generalizations appear warranted by these reports.

---

[1] Wheeler, Raymond H., The Synaesthesia of a Blind Subject, *University of Oregon Publications*, Vol. I, No. 5, May, 1920, p. 32.

1. *A low tone.* Thirty-four persons associate low tones with dark colors or grays. Fourteen associate no color or brightness with low tones. One associates with medium gray, and another with "rich colors."

2. *A tone of medium pitch.* There is no uniformity of color here, but a generally uniform medium brightness in 26 cases. The remaining 24 discover no such association.

3. *A high tone.* This is associated in all but 14 cases with bright colors or light grays. There is a wide variety of color associations.

4. *The key of C.* Twenty persons report a color or brightness associated with this key, eight of these associating it with blue. The dominant association, however, seems to be with medium brightness, which may be attributed to the commonness of this key in music and the medium range of compositions.

5. *The key of A.* Only nine persons mention color assocations with this key, and these vary from blue to yellow.

6. *Key of D minor.* Fifteen persons associate this key with dull or soft colors, as pale green, lavender, or violet, and with grays.

7. *Key of A flat.* Only ten report visual associations and these show no uniformity.

8. *The piano.* Twenty-four persons associate the piano with colors or brightnesses, but there is no observable uniformity. The piano may perhaps represent a great variety of imagery according to key, pitch, etc.

9. *The violin.* Only fifteen have no color or brightness imagery associated with the violin. To nearly all the remaining 35 it suggests delicate colors, predominantly blue, violet, and related colors. In one case there was a traceable association between brightness or delicate color and the light dresses of violinists heard at Chautauqua.

10. *The pipe organ.* Nineteen have no color or brightness association with the organ. Most of the remainder associate it with dark values, brown being the one most frequently mentioned.

11. *The clarinet* suggests to 19 persons colors of high brightness value, but has no such association for the remaining 31.

12. *The trombone.* Fourteen reports associate this instrument with dark colors or yellow, with brown predominant.

13. *The cello.* In 19 reports the browns and reds predominate, perhaps due, in some degree, to association with the color of the instrument.

14. *The snare drum.* Twenty persons report an imagery usually of bright colors or gray. In four cases there is a mixture of colors or brightnesses.

15. *The bass drum.* Twenty-eight persons report an imagery chiefly of low brightness values.

16. *The flute.* To twenty-four persons the flute suggests colors of high brightness values, or white.

17. *Dvorak's Humoreske.* A large proportion of the students have color association with this familiar composition, but there is no uniformity. Some appear to distinguish different color or brightness imagery corresponding to the very distinct moods of the piece.

18. *Mendelssohn's Spring Song* is rather naturally associated with spring colors, greens, pinks, etc., in 33 cases.

19. *Chopin's Funeral March* is associated by 32 persons with dark shades and such colors as blue, violet, and lavender. The use of lavender, as well as black and grey, in funeral decoration should be noted.

20. *The Star Spangled Banner* could with difficulty be dissociated from the colors of the flag itself. However, 26 persons associate it with single colors, chiefly red and yellow.

21. *Annie Laurie* suggests to 27 people soft and relatively dull colors, 17 of them mentioning blue or green.

22. *The Soldiers' Chorus from Faust* is associated in the experience of 22 persons with bright colors, red and blue being predominant.

23. *The Pilgrims' Chorus from Tannhauser* suggests medium or dark colors or brightnesses to 17 people.

24. *A lively march* suggests red to 19 persons and other colors to 10 more.

25. *A dreamy waltz* has color or brightness associations for 29 persons, chiefly the softer colors of relatively high brightness, as light blue, pale yellow, pale pink.

It is clear that much, perhaps most, of the visual imagery aroused by musical selections or even by given musical instruments is derived from particular experiences and associated circumstances. Such imagery is apt to be very complicated and shifting, and hence anything like Scriabin's " colored symphony " involves immense, perhaps insuperable, difficulties.

There are, however, certain observable tendencies to uniformity, whatever their origin may have been. A high tone or a high pitched instrument is generally associated with bright colors or lighter grays, while a low tone or a commonly low pitched instrument suggests dark color and brightness values. There appears also to be a relation between the timbre of an instrument and the richness of color associated with its tone. Although there is a wide variation of color-imagery, it is probable that the common term "tone-color" is not a mere figure of speech based upon analogy. A tone which approaches a "pure" tone is described by musicians as "white or colorless." It is probably not a mere musician's whim that commonly associates complexity of tone with the visual richness of color.

A COMPARISON OF TWO METHODS OF GIVING THE NUMBER SERIES COMPLETION TEST.

JOHN E. ANDERSON, Yale University

Incidental to an investigation upon intelligence tests with Yale Freshmen a comparison was made of the cross-out method as developed by Pressey[1] and the completion method as used in the army of giving Alpha Test 6, the number series completion test.

The cross-out test was constructed by taking Test 6 Alpha Form 9, completing each series and inserting a number not in the series in each row in such a way that at first glance it might appear part of the series. The subjects were given instructions to cross out the number in each row which did not belong in the series. The Test 6 Alpha instructions and examples were paraphrased to fit the cross-out method, being kept as nearly as possible in the original form.

A preliminary examination of 14 subjects demonstrated that the Alpha Test 6 time limit of 3 minutes was much too brief for the cross-out method and resulted in the selection of 7 minutes as the time limit.

Eighty-five subjects were used, obtained by requesting the men of three divisions of a required subject to report for a second examination which consisted of the cross-out test among several others. These men had previously taken Alpha Form 6 with the entire group of Yale College Freshmen.[2] An interval of seven months elapsed between the two examinations.

The results are as follows:

	Time Limit	Range	Mean	Standard Deviation
Completion Method............	3 mins.	5–20	13.5	2.7
Cross-Out Method.............	7 mins.	4–20	12.7	3.4

Although the time limits differ to a considerable extent, the two methods of giving the test result in practically the same range, mean, and standard deviation.

---

[1] Pressey, S. L. and Pressey, L. W. *Journal of Applied Psychology,* Vol. III, 1919, p 138-150.

[2] Anderson, J. E., *School and Society,* Vol. XI, 1920, pp. 417-20.

Correlation coefficients were obtained by the product moment method between the average college standing of the men for the entire Freshman year's work and the test scores. These are as follows:

	r.	P. E.
College standing and Cross-out	+.25	±.07
College standing and Completion	+.04	±.07
Completion and Cross-out	+.44	±.06
Alpha score and College Standing	+.25	±.07
Alpha score (with cross-out scores substituted for completion scores) and College Standing	+.34	±.06

Unfortunately I have no measure of the reliability of either of the methods because the tests were not repeated. But the coefficient of correlation between the completion method and the cross-out method of giving the test is low enough to indicate that the change in the method of administration results in a distinct test, different from the original, a conclusion borne out by the other coefficients. While the cross-out method correlates to some extent with college standing the completion method does not. The effect of substituting cross-out scores for completion scores in Alpha is shown by the increase in the resultant coefficient for total score with college standing from .25 to .34.

The evidence from 85 cases of Yale Freshmen is that the cross-out method of giving the number series completion test is superior to the regular completion method as measured by the criterion of correlation with college standing.

THE APPLIED PSYCHOLOGY OF NAMES

By A. A. Roback, Harvard University,
Assisted by Marie Groetzinger

1. *Nature of the Problem*

Though the field of memory has been traversed carefully by so many painstaking investigators since the days of Ebbinghaus, new problems have come up through channels which until recently had not been amenable to the scrutiny of general psychology. The practical demands of applied psychology opened the way for many a new problem; and as is frequently the case, the results gained in the practical sphere may offer some suggestions as regards the theoretical aspects of memory.

In our present commercial system, it is becoming increasingly common to see firms consisting of two or more partners, which often necessitates their names being arranged in a certain combination. Usually what determines the precedence of one name over another is the seniority of one of the partners. But it is not always that a firm can claim a senior partner. Suppose two men agreed to enter into a business relationship with the same capital and with equal privileges and liabilities. The question before us then is: is there any general rule that can be applied as to the way the combination of their names should appear in advertisements, on letter-heads *or in any form of publicity? If Smith and Stanft were to start* business together, would it be more advantageous to have the names stand as they were just mentioned, or would it be more advisable to reverse the order and call the firm Stanft and Smith? It will readily be perceived that the issue is one between familiarity and unfamiliarity, since in most situations of that sort, one name is not so familiar as the other. There are, to be sure, other differences between names, but that of familiarity or unfamiliarity seems to be the most comprehensive, embracing even the length element of the name. We often hear people say " Oh, I can't think of the name. It's such an *uncommon* name." The memory lapse is thus directly ascribed to unfamiliarity.

Were it merely a matter of settling the question whether familiar names are more easily remembered than unfamiliar ones, experiments would scarely be neccessary. The principle of association implies this result unmistakably, and in practically every memory investigation there are indications of the greater mnemonic value in the case of the more familiar material.[1]

In confining our problem to the study of combinations consisting of familiar and unfamiliar names, we have a definite task to perform, since no information is as yet available on this phase of memory, and in default of such knowledge, business men who are confronted with the situation referred to above, will decide the question *a priori* or perhaps by tossing a coin—methods which are not permissible in either science or business.

In the course of the investigation, a new item presented itself for further inquiry, viz, the fact that repetition offers a greater gain in effective value for unfamiliar names than it does for familiar ones. This apparent result prompted another set of experiments in order to ascertain whether the gain is substantial enough to make up for mnemonic difficulty going with unfamiliarity. For theoretical purposes, the bare result would be important in itself, but the practical side of the problem thus becomes encumbered by another situational factor, so that if it is established that the one type of combination has an advantage over the other, the question may be asked: after how many exposures? It may be that after the first repetition, the name "Cui" has a better chance for reproduction than such a common name as Smith or Jones but that subsequently, either on account of the many associative inhibitions occurring with the latter or else because of the disproportionate distribution of attention, the name "Cui" would actually be the one to be recalled more easily in connection with a given firm.[2]

That the applied psychology of names is no minor factor in business can be inferred from the fact that so many aliens and Americans of foreign descent find it expedient either to change their names entirely or else modify them in some way as to conform to the general vogue of names: the "owitz" or "sky" termination in Slavic names is frequently dropped; the *Umlaut* and duplication of letters so common in German names are often done away with, and in general there is a tendency to adapt the name to suit the native taste. Most foreigners who enter an application for such changes state explicitly that it is a business measure; and they seem to be satisfied with the results.

---

[1] It must be noted that a common name and a familiar name are not interchangeable terms. An uncommon name may become familiar through some political event or because of the rise to popularity of some theatre (movie) star. Such accidents will of course, to a considerable extent, affect the degree of familiarity or unfamiliarity of a name without appreciably reducing its uncommonness. Contingencies of this sort cannot enter into our problem, though they doubtlessly have a decided bearing on our memory function.

[2] The repression of certain names on Freudian principles, if at all operative, is altogether too individual an occurrence to render it worthy of consideration in this connection. Should we accept the extreme view that every memory lapse is symptomatic of a repression, the whole of the psychology of memory would reduce to a speculative psychoanalysis barren of the least predictability.

The subject is not devoid of its ethical and legal aspects. It may be questioned whether one with a true sense of family pride would assume another name. Still more may it be disputed whether it is commendable to give up from practical motives the most conspicuous ancestral heritage. When we note the persistent adherence of the Savages, the Cowards, the Hoggs, the Butchers, the Swindles and the Coffins to their ill-sounding, but at the same time significant, patronymics, we cannot but admire their loyalty to their predecessors as well as their independence in maintaining their familial identity unbroken.

Happily our investigation is not concerned with moral issues nor is it called upon to consider personal whims or tastes or the domineering influence of one partner over the other in the matter of name combinations. We shall assume, on our side, that the only factor which is to determine the arrangement of the parnters' names is to be that of promoting their several business interests; and within that compass, our problems lends itself to a clear-cut treatment.

The practice of adopting impersonal trade names, such as is a necessity with corporations, will at once suggest itself as a welcome compromise whenever there is a question that is likely to become serious as to the priority of a particular name. There are indeed numerous lines of business which favor the impersonal trade name. A garage, for obvious reasons, will almost invariably be known by its locality. Similarly we are accustomed to such signs as The City Coal Co., The National Tea Co., The United States Tire Co., The New York Paper Co., The College Bindery, etc. etc. Manufacturers of novelties and even standard articles are especially apt to link their wares in the minds of the public with an impersonal symbol suggestive of (in many cases inappropriate to) the particular article. The marketable value of this symbol as a brand or trademark or label has been so much emphasized in books on advertising that the mere mention of it here as bearing a tangential relation to our problem is sufficient.

It will readily be conceded that the personal name is still in vogue among most firms and particularly in certain lines of merchandise, where the adoption of a patented trademark does not prevent the merchants from identifying themselves in connection with the wares.

2, *Method of Procedure*

The experiments to be reported were carried on in the Psychological Laboratory of the University of Pittsburgh during both semesters of 1917-18. The subjects were two instructors in the department of psychology, and one woman student, with considerable training in psychology.

Apparatus. The apparatus consisted of a kymograph and two screens. The larger screen was made of black card board with an opening 3½ × 2 inches. The smaller screen was made of tin, painted black, and curved to fit the kymograph drum. The opening in the second screen (2½ × ¼ inches,) was large enough to expose one stimulus at a time from the list revolving on the kymograph drum.

Material. Familiar and unfamiliar names were used in single columns and in combinations of (a) familiar with unfamiliar, and (b) unfamiliar with familiar. The names were arranged and grouped as follows:

1.	20	Monosyllabic	familiar
	20	monosyllabic	unfamiliar
2.	20	dissyllabic	familiar
	20	dissyllabic	unfamiliar
3.	15	monosyllabic	familiar with unfamiliar
	15	monosyllabic	unfamiliar with familiar
4.	15	dissyllabic	familiar with unfamiliar
	15	dissyllabic	unfamiliar with familiar
5.	10	trisyllabic	familiar with unfamiliar
	10	trisyllabic	unfamiliar with familiar

Examples of familiar names—Brown, Dixon, Foster, Wallace, Matthews, etc.

Examples of unfamiliar—Farquhar, Martinez, Swoboda, Enza, etc.

Six, or eight lists of names were given at each sitting, the number varying on account of the difference in the length and the complexity of the lists.

At each sitting the lists of the single names were given first, followed by the lists of combination, i. e. beginning with the simpler names and then going to the more complex combination. The subject read aloud each stimulus as he saw it. When the list was complete he would write down as many of the combinations and single components of the combinations as he could remember.

During the first semester one review of the list was given, but during the second semester two reviews were given of each list so as to test whether there was any tendency to remember the unfamiliar names better after repetition than the familiar names. In this way the subject at each sitting got some lists that were entirely new while others he merely had to review.

It was rather difficult to get names that all the subjects would regard as either decidedly familiar or unfamiliar. The aim of the experimenter was of course to eliminate all indefinite cases, though this was not possible at all times. Local differences naturally cannot but enter into the determination of the degree of familiarity. Thus a name like Shanahan would be considered familiar in Pittsburgh though elsewhere it might not be so regarded.

It took fifty seconds to expose a list of single names and one minute to expose a list of paired names. All of the names were typewritten on long sheets of paper so as to fit around the kymograph drum.

One of the subjects required a double exposure of each list in order to reproduce at least several of the names.

The first day the series of single familiar names was exposed before the unfamiliar names. The same procedure was gone through with the names in combination—the familiar-unfamiliar combination being given before the unfamiliar-familiar. At the second sitting, however, this order was reversed and the unfamiliar names were exposed first. This alternating method was used throughout the experiment in order to eliminate the well known factor of *primacy*, according to which the first elements are remembered better than the succeeding ones.

Key to Tables

The tables for the monosyllabic, dissyllabic and trisyllabic combinations give in separate columns the total percentage both for the combination as well as for the separate single components i. e. either the familiar or the unfamiliar element remembered.

After repetition there is always a gain as may easily be expected, but in one or two cases, curiously enough, a loss is shown, in which case the numbers are preceded by a minus sign.

There were no third trials given during the first semester, which explains why the results for the first and second trials are based on so many more combinations than the results for the third trials.

During the first semester two weeks intervened between successive trials, but during the second semester the interval was reduced to one week.

TABLE No. 1

Total percent. of names remembered out of 360 single monosyllabic and 120 single dissyllabic names.

	No. of Stimuli	First Trial		Second Trial	
		Familiar	Unfamiliar	Familiar	Unfamiliar
Total percent. of names remembered from the monosyllabic single series........	360	51%	35%	54%	41%
Total percent. for the dissyllabic single series..........	120	42%	24%	42%	26%
Total percent. for the monosyllabic and dissyllabic single series.............	480	46.5%	29.5%	48%	33.5%

Note on Table 1.

This table shows the total percent. remembered out of 360 single monosyllabic and 120 dissyllabic names. As would be expected, during the first trial there were more familiar names remembered than unfamiliar; and in the second trial there was the natural increase in every case except the dissyllabic familiar, due perhaps to distraction on the part of one of the subjects in this particular trial. These results also reveal the fact that the unfamiliar names show a greater percentile gain in the second trial over the first than do the familiar ones. In the monosyllabic names there was a gain of 6% in the unfamiliar while in the familiar series there was a gain of only 3%. Furthermore, the gain for the unfamiliar dissyllabic names was 2% while the familiar dissyllabic names showed no gain. This general tendency appears likewise in the total percentage of the combined monosyllabic and dissyllabic lists.

TABLE No. 2

Total percent. of names remembered out of 90 *monosyllabic* names in combination.

		First Trial						Second Trial					
		Familiar—Unfamiliar			Unfamiliar—Familiar			Familiar—Unfamiliar			Unfamiliar—Familiar		
	No. of Stimuli	C	F	U	C	F	U	C	F	U	C	F	U
Total percent. for Monosyllabic Combination..........	90	22%	4.4%	8%	21%	9%	4.4%	30%	5.5%	16%	22%	10%	5.5%
Gain of second trial over first trial..........								8%	1.1%	2%	1%	1%	1.1%

Note on Table 2

This table gives the total percent of names remembered in *combinations* of monosyllabic familiar and unfamiliar names such as Smith and Knaus, Tsolt and Hughes. The gain of the second trial over the first in these combinations is also given.

In both the first and second trials there were more familiar-unfamiliar combinations remembered than unfamiliar-familiar ones. As is shown in the gain of the second trial over the first, the unfamiliar half of the combination benefits more by the repetition than the familiar half.

TABLE No. 3

Total percent. of names remembered after the first and second trials of the combination of 540 dissyllabic names.

	First Trial						
	No. of Stimuli	Familiar-Unfamiliar			Unfamiliar-Familiar		
		C	F	U	C	F	U
Total % for dissyllabic combination............	540	17.3%	10.4%	4.4%	14.9%	14.7%	4.9%
Gain or loss of second trial over first trial........							

TABLE No. 3—*Continued*

		Second Trial					
	No. of *Stimuli*	Familiar-*Unfamiliar*			Unfamiliar-*Familiar*		
		C	F	U	C	F	U
Total % for dissyllabic combination.	540	10.4%	8.9%	10.7%	17.1%	15.3%	7.8%
Gain or loss of second trial over first trial............		6.9%	1.5%	6.3%	2.2%	.6%	2.9%

Note on Table 3.

This table gives the total percent. of names remembered after the first and second trials of the dissyllabic combinations of 540 names, and also the gain or loss in the second trial over the first. There was a gain in every case, which naturally was to be expected, except in the case of the whole combination in the familiar-unfamiliar list which shows a loss of 6.9%. This loss however, is to be considered together with the gain in the familiar and unfamiliar separate names which formed part of the combination, so that in reality the loss was much slighter; and if we give unit values to the combination and half values to each separate component (of the combination) then the loss for that series would be only 3%. This loss, which seems peculiar, may be explained by the physical condition of one subject—at least there is a note to that effect in the protocol.

In the first trial there is a greater percent. of names remembered in the familiar-unfamiliar combination than in the unfamiliar-familiar. However, this is reversed in the second trial where there were 17.3% remembered in the unfamiliar-familiar combination over against only 10.4% in the familiar-unfamiliar combination. This result is not incompatible with the first result, i. e. that the familiar-unfamiliar combination is more favorable for remembering than the unfamiliar-familiar, but is rather in accord with what has afterwards been found to be the case in regard to the effect of repetition on the familiar names as compared with the unfamiliar names.

This table also shows a greater gain in the second trial of the unfamiliar over the familiar, since in both types of combinations (the familiar-unfamiliar and the unfamiliar-familiar) the unfamiliar components showed in the second trial a much greater gain than the familiar components, viz. 6.3% and 2.9% as against 1.5% and .6%.

TABLE No. 4

Total percent. of names remembered after the first and second trials of the combinations of 60 trisyllabic names.

First Trial

	No. of Stimuli	Familiar-Unfamiliar			Unfamiliar-Familiar		
		C	F	U	C	F	U
Total percent. for the trisyllabic Combinations.	60	11.6%	20%	10%	20%	13.3%	15%
Total percent. for the trisyllabic Combinations Gain of second trial over first, using only two subjects..............	40	10%	25%	5%	12.5%	20%	15%

Second Trial

	No. of Stimuli	Familiar-Unfamiliar			Unfamiliar-Familiar		
		C	F	U	C	F	U
Total percent. for the trisyllabic combinations.............	60	25%	13.3%	11.6%	21.6%	23.3%	15%
Gain of second trial over first trial.....		13.4%	-6.8%	1.6%	1.6%	10%	0
Total percent. for the trisyllabic combinations of two subjects.............	40	20%	20%	7.5%	20%	30%	12%
Gain of second trial over first, using only two subjects..		10%	-.5%	2.5%	7.5%	10%	-.3%

Note on Table 4.

This table giving the total percent. of names remembered out of a series of 60 trisyllabic names in combination such as Richardson and Boyarjan was inserted only for the sake of completeness. As a matter of fact its value may be questioned not only because the number of stimuli was insufficient to warrant any conclusion but also for the reason that some of these names which were thought to be unfamiliar by the experimenter were *really familiar* to one of the subjects. This is the only series in which there is no correlation apparent between the arrangement of the combination and its memory value, even after eliminating the results of the one subject whose acquaintance with a certain class of names proved to be a disturbing factor.

TABLE No. 5

Total percent. of names remembered after the *third trial* of the combination of 180 dissyllabic names.

	No. of Stimuli	Familiar-Unfamiliar			Unfamiliar-Familiar		
		C	F	U	C	F	U
Total percent. for the third trial dissyllabic combinations.	180	21.6%	8.3%	11.1%	18.7%	17.2%	13.8%
Gain or loss of *of* third trial over first trial..	180	3.3%	—2.1%	6.7%	3.8%	2.5%	8.9%
Gain or loss *of* third trial over second........	180	10.2%	— .6%	.4%	8.3%	1.9%	6%

Note on Table 5.

This table gives the total percent. for the third trial of the dissyllabic combination of 180 names and also the gain or loss over the first and second trials.

These results show the natural increase over the first and second trials. The only place showing a loss is in the familiar column of the familiar-unfamiliar combination. This loss however, is very slight and is made up for by the gain in the other columns. Giving unit values to the complete combination and half values to the separate components, there is a total gain of 10.1%.

In every case the unfamiliar names show a greater gain upon repetition than the familiar ones.

Note on Table 6.

This table gives the total percent. of names remembered after the third trial of the trisyllabic combinations of 60 names, and also the gain or loss of the third trial over the first and second trials. The results are the same as in table 5

TABLE No. 6

Total percent. of names remembered after the third trial of the combination of 60 trisyllabic names.

		Familiar-Unfamiliar			Unfamiliar-Familiar		
	No. of Stimuli	C	F	U	C	F	U
Total percent. for third trial of trisyllabic combinations............	60	31.6%	15%	21.6%	25%	21.6%	28.3%
Gain or loss of third trial over first....	60	20%	—5%	11.6%	5%	8.3%	13.3%
Gain or loss of third trial over second..	60	6.6%	1.7%	5%	4.4%	—1.7%	13.3%

in that the familiar-unfamiliar combination is favored to begin with, but after repetition, the unfamiliar components of both the familiar-unfamiliar and unfamiliar-familiar combinations show a greater mnemonic gain than the familiar components.

Note on Table 7.

This table gives the total percent. for all the series of combined names, thus showing the relation between the familiar-unfamiliar and unfamiliar-familiar combinations. The third trial concerns only the combination of 180 dissyllabic names and 60 trisyllabic names, since no third trials were given for the monosyllabic combinations. In the first and second trials

TABLE No. 7

Total percent. of names remembered for *all the series* of combined names.

First Trial							
	No. of Stimuli	Familiar–Unfamiliar			Unfamiliar–Familiar		
		C	F	U	C	F	U
Total percent. for all series of combinations............	690	16.9%	11.6%	7.5%	18.9%	12.3%	8.1%
Total percent. counting combinations as 1, and single components as ½.....	690	26.4%			29.05%		

TABLE No. 7—*Continued*

Second Trial

	No. of Stimuli	Familiar-Unfamiliar			Unfamiliar-Familiar		
		C	F	U	C	F	U
Total percent. for all series of combinations..............	690	21.8%	9.2%	10.4%	20.2%	16.2%	9.4%
Total percent. counting combinations as 1, and single components as ½..	690	31.7%			33%		

Third Trial

	No. of Stimuli	Familiar-Unfamiliar			Unfamiliar-Familiar		
		C	F	U	C	F	U
Total percent. for all series of combinations............	240	26.6%	11.6%	16.3%	21.8%	19.4%	21%
Total per cent. counting combinations as 1, and single components as ½............	240	40.59%			42%		

the results are taken from 690 combinations of monosyllabic, dissyllabic, and trisyllabic names, but in the third trial the results concern only the combination of 240 dissyllabic and trisyllabic names. These results might seem to reveal a slight preference for the familiar-unfamiliar combination over the unfamiliar-familiar. However, by giving unit values to the whole combination and half values to each of the familiar and unfamiliar components, it becomes evident that the unfamiliar-familiar combination is favored, due to the greater number of single unfamiliar components remembered as compared with the single familiar components.

TABLE No. 8

Total *gain or loss* for all the series of combined names.

	Familiar–Unfamiliar			Unfamiliar–Familiar		
	C	F	U	C	F	U
Total gain or loss of second trial of combination series over first trial...........	4.9%	**−2.4%**	3.2%	1.3%	3.9%	1.3%
Total gain of third trial of combination series over first trial...............	9.7%	.05%	8.85%	2.95%	7.1%	12.95%
Total gain of third trial of combination over second trial..................	4.8%	2.45%	5.65%	1.65%	3.2%	10.65%

Note on Table 8.

This table shows the relation between the gain or loss of the second trial over the first, the third trial over the first trial and the third trial over the second; and the relation between the gain in the unfamiliar as over against the familiar. In every case the gain after repetition is greater in the unfamiliar than in the familiar column.

Conclusion

The above results tend to prove (1) that a combination of names possesses a greater immediate memory value if the more familiar component of the combination appears last, and the less familiar first; (2) Other things equal, the unfamiliar element shows a greater memory gain than the familiar component—which sugests that in time, that is to say, after a large number of repetitions, the former will become a more effective stimulus than the latter.

INTELLIGENCE TESTS AND ACADEMIC STANDING

EDNA GORDON and H. J. BAKER, University of Michigan
Communicated by W. B. PILLSBURY, Professor of Psychology,
University of Michigan

According to the survey made by Miss Caldwell[1] there is not much evidence for a close correlation between the results of mental tests and college grades, the range in coefficients of correlation between these two factors, including Miss Caldwell's own experiment is from .09 to .44.

The following report adds no conclusive evidence to that already obtained by others, but the results given here correspond somewhat to those obtained by Miss Caldwell.

As a part of the course in the study of mental tests, forty-four students in an elementary course in Psychology were given the Stanford Revision of the Binet-Intelligence Tests XIV, XVI (Average Adult), XVIII (Superior Adult).

The majority of the students were sophomores, the remainder upperclassmen in the University. The average chronological age was twenty-one years, three months; mental age, approximately seventeen years, ten months. The range in intelligent quotients was 91 to 119, average 111.

The grades of these students in the various subjects pursued were obtained from the Registrar's office. Using the ratio of hours to honor points, (a C grade in a four-hour course gives four points, B grade, eight points, A grade, twelve points, D grade no points) a general ratio was obtained for each student in all subjects, and in each group of subjects as follows: Languages, Rhetoric and English, History, Philosophy, Political Science, Political Economy and Sociology, Oratory and Fine Arts, Mathematics, Physics and Chemistry, Other Sciences.

The coefficient of correlation, Pearson formula, was obtained between the I.Q. and the general ratio (hour-honor points) for the group in all subjects; and for each group of subjects as shown in the following table:

---

[1] Helen Hubbert Caldwell, "Adult Tests of Stanford Revision Applied to College Students." *Jour. Educational Psychology,* Dec. 1919.

I. Q. Correlated With		Probable Error
General Ratio	.446	Less than r/5
Political Ec.—Soc.	.548	" " r/5
Other Sciences	.537	" " r/5
Rhetoric—English	.481	" " r/5
Mathematics	.450	" " r/4
Oratory—Fine Arts	.354	" " r/2
Physics—Chemistry	.334	" " r/3
Languages	.229	" " r/3
Philosophy	.096	Greater than r
History	.047	" "
Political Science	.017	" "

It will be noted that the probable error is so large in the last three cases that no significance can be attached to the results; in fact, in all the coefficients of correlation given in the above table, it must be borne in mind that the number of students and the number of hours represented in the different subjects were widely variable. Therefore no attempt will be made to draw conclusions save the very general one, that on the whole there is a slight positive correlation between the results of this intelligence test and academic standings, and that further and more extended investigation is needed before one can definitely state what the relationship is between these two factors.

In addition to the college students, twenty-four students in a short term course in Public Health Nursing were given tests. The coefficient of correlation between the intelligence quotients and the grades in psychology is .50. The following summer a second group of twelve nurses were tested by another class and a coefficient of correlation of .21 obtained between intelligence quotients and grades in psychology. A third group of eighteen nurses were tested during the first semester of this year, and a coefficient of correlation of —.187 was obtained. Owing to some difficulty and misunderstanding in the administrative offices, we were unable to obtain the records of the nurses in other courses.

The last two correlations we consider of little value, because of the small number of cases, and although the first coefficient, .50, does seem significant, we cannot place as much confidence in the results of the tests given the nurses as in the results obtained from testing the regular students. In first place, the nurses were of widely varying education and experience, some being middle aged, with no high-school training; others were high-school graduates and had but recently finished their nurse's training. The regular college students, on the other hand, were practically of the same age and experience. (Then too, the inexperience of the examiners may account in some measure for the apparent discrepancy in results.) The college students were all tested by three persons who had previous experience in giving tests.

A DECISION STUDY OF 150 YOUNG MEN AND WOMEN

By SYBIL MARSH GIBSON

1. *Description of the Test*

The following study was undertaken in the Harvard Psychological Laboratory in 1915 with the purpose of standardizing a decision test for use with psychopathic and defective cases. In the search for norms with which to compare the performances of such cases, comparative data from normal individuals, men and women, were obtained, and these are presented here as having interest in themselves.

The aim was to have a series of decisions of graded difficulty and of uniform difficulty within a given grade so that some idea might be obtained of the characteristic time reaction of the individual to decision situations of a given difficulty, to such situations as compared with situations of less and greater difficulty, the constancy of such time reactions, and the correctness of the resulting decisions.

Professor Muensterberg's so-called judgment test[1] was chosen for a starting point. The test from which the present data were obtained is, however, much changed from the Muensterberg test. The object of the changes has been to make the test more adapted for bringing out the decision reactions mentioned above.

The capital letters E, I, Y, and K were selected after experimentation as having more uniform appeal to attention than the Muensterberg letters, A, E, O, and U, and unlike the Muensterberg test, no additional letters were used. These four letters are arranged in five rows of ten letters each on twenty-four cards three and a half inches long by two and three-fourths inches wide, the letters occupying a block at the top of the card. Only four arrangements of letters were used throughout the series of twenty-four cards, giving four situations of graded difficulty as follows: grade I, the most easy of decision, in which, the predominating letter appeared twenty-six times, and the three non-predominating letters eight times each; grade II, in which the predominating letter appeared twenty-three times, and the non-predominating letters nine times each; grade III with the predominating letter twenty times, and the non-predominating ten times each; and grade IV, the most difficult, in which the predominating letter appeared seventeen times and the non-predominating eleven times each. These four arrangements of letters were selected after experimentation, a general mixing of the letters being sought instead of "catchy" arrangements in which the predominating letters were arranged on the edge of the field, etc., that the basis for difficulty might be numerical so far as possible. The aim in the selection of total number of letters, number per line, number of lines and arrangement within the lines was to obtain four grades of difficulty which would, on the average, take for decision times from two seconds for the easiest grade to ten or fifteen seconds for the most difficult. The averages of the groups included in this study approximate these figures.

---

[1] Hugo Muensterberg, *Psychology and Industrial Efficiency* pp. 83-96, and J. W. Bridges, *Psychological Review Monographs*, Aug. 1914, pp. 51-55.

Four cards make up grade I and IV respectively, and eight cards grades II and III respectively, different predominating letters, merely, being substituted in the four arrangements. While it was desired that all cards within a given grade should be identical with the exception of the different predominating letter, this is obviously impossible, since, for example, on a K card we might have a K between an I and an E, but on an E card we cannot have an E between and I and an E, but must substitute some other non-predominating letter for the second E. All that was possible was to have the predominating letters occupy the same spatial position on the cards, with definite spatial relations to non-predominating letters, though not always to the same non-predominating letters.

Since there are eight cards in both grades II and III, it is obvious that not only does the same arrangement appear eight times in each of these grades, but that in four cases in each, the identical card appears twice. Preliminary experimentation and experience with the test both give little ground for objection to this on the score of greater facility in decision on repeated arrangements, or repeated cards. Apparently a card of the same arrangement appears very different when a different predominating letter is substituted, and the appearance of any given card is quickly forgotten after a new card has been presented. Sometimes even two cards exactly alike can be presented to a normal subject in sequence without his noticing that they are identical. However, this extreme situation never occurs in the present test.

The manner of giving the present test is also very different from that of the original Muensterberg test. The cards are given in a definite sequence to each subject so that, unlike the Muensterberg test in which the cards were shuffled before beginning, each subject takes exactly the same test. This manner of giving the cards likewise allows the timing of decisions on separate cards. The principle in devising the sequence in which the cards should be given was to have no very difficult cards at the beginning before the subject was accustomed to the test or to discourage him at the task before him, no very difficult cards at the very end when there might be some measure of fatigue, no identical cards in succession, and no two difficult cards in succession.

The directions given the subject before he undertook the test, unlike the Muensterberg test, emphasized accuracy and not speed. It was desired to reproduce, as nearly as possible. the condition of decisions of real life, of which quick decisions form merely one group. There is to be sure always some time element in actual decisions, and a suggestion of this was brought into the directions by telling the subject that while he was not limited as to time, his time was being recorded for every card.

The decision and time of decision on each card was then recorded by the examiner. The time was recorded merely in seconds by an ordinary second hand. Exact laboratory methods were not desired for a test that was planned for free use outside the laboratory in connection with routine scales. However, since separate cards are timed and time decision on grade I was often one or two seconds, it is felt that the method is not fine enough, and that a stop watch should have been used to measure fractions of a second. However, though the rougher method was used in obtaining the data presented here, the differences in the individual records and between the two groups are so striking as not to be attributable to the roughness of the method, and so seem to be of value in spite of it. Accuracy records are, of course, not affected by this. The lack of introspective data from the subjects is also to be regretted.

In evaluating the individual records, the following measures were used: total time taken for the test, average time spent on a card of each grade, porportional variation of the decision times of the individual on each grade $\left(\frac{\text{M. V.}}{\text{Mean}}\right)$, accuracy on each grade, and the six ratios between the average four grade times to show how much time the individual spent on harder decisions as compared with easier ones. The use of the measure of proportional variability involved much labor which seemed out of proportion to the results obtained from it. While some such measure was needed to show how constantly an individual spent a given amount of time on problems of the same degree of difficulty, only large variations seem to be very important since all of the normal individuals showed considerable variability, and such high variability can be fairly easily discovered on glancing through the record. The measures of total time, average grade times and grade accuracy are, of course, essential and easily obtainable. The use of the six ratios is less readily obtained and makes the test less easy to evaluate. In comparing a single record with the norm for the group, the six ratios can be fairly accurately estimated at a glance, the detailed figuring being necessary only in comparing groups.

2. *Comparison of Men and Women.*

The test was given to 150 men and women, college students. They were, for the most part, undergraduates, with a few graduate students. They were not chosen for any special qualities or abilities, but constituted a general mixed group which volunteered for the test. The average age of the two groups was practically the same, that of the 75 men being 20.24 years, and that of the 75 women 20.99. The groups were also both fairly homogeneous in age, the mean variation from the average ages of the group being .99 of a year with the men, and .77 of a year with the women.

Both groups were being tested at the same time, and by the same examiner. The examining was done before the make-up of the classes was changed by the war.

Comparing the performances of the two groups for decision time, we find that on each grade each group took more time on the average than on the preceding grade. Each group, then, spent on the average more time for decision on harder situations than on the easier ones. The average times for the two groups were as follows:

	Total Time	Grade I	Grade II	Grade III	Grade IV
Men............	2′ 57″	2.76″	3.99″	6.10″	9.10″
Women..........	2′ 8″	3.59″	5.05″	8.18″	14.34″

It is also evident from these figures that the women took more time, on the average, on each grade than the men, with the difference slightly more marked on the more difficult grades, the relation of women's time to men's time being as follows:

Grade I	1.30 : 1
Grade II	1.26 : 1
Grade III	1.34 : 1
Grade IV	1.54 : 1

The limits of variation with both groups were wider on each succeeding grade, with the reaction times of both groups least homogeneous on grade IV. The women group showed wider limits of variation on each grade than the men group.

Comparing the two groups for accuracy, the women had fewer mistakes on all four grades than the men.

On grade I, only one mistake occurred among the women, and seven among the men, an average of .013 and .09 respectively. The seven mistakes of the men were distributed among five individuals, three making one mistake apiece, and two men two each.

On grade II the men made thirty-two mistakes and the women sixteen, an average of .43 and .21 respectively. The sixteen mistakes of the women were distributed among thirteen individuals, ten making one mistake apiece and three making two apiece. The thirty-two mistakes of the men were distributed among fifteen individuals, six making one mistake apiece, three making two apiece, five making three apiece, and one making five. There were then approximately the same number of individuals in each group making mistakes in the grade, but there were more individuals among the men making more than one mistake apiece.

On grade III, the men had ninety-nine mistakes, and the women eighty-three, an average of 1.32 and 1.10 respectively. The eighty-three mistakes of the women were distributed among forty-two individuals, seventeen having one apiece, twelve two apiece, seven three apiece, four four apiece and one having five. The ninety-nine mistakes of the men were distributed among fifty-four individuals, twenty-five having one apiece, seventeen two apiece, eight three apiece, and four four apiece. On this grade, then, there were more individuals making mistakes among the men than among the women, but not more men making a large number of mistakes.

On grade IV, the men made one hundred and sixty-four mistakes, and the women one hundred and thirty-five, an average of 2.18 and 1.77 respectively. The mistakes were distributed through practically the entire groups (seventy-two men and seventy-one women) as follows:

	Women	Men
1 mistake	26	17
2 "	27	26
3 "	17	21
4 " (all)	1	8

More men than women, then, made three or four mistakes and fewer only one or two mistakes.

While the women, then, were more accurate than the men on all grades, this difference is more marked on the two easier grades.

In order to compare the accuracy per individual for the whole test among the two groups, a kind of arbitrary point scaling was adopted. A mistake in grade IV was counted one, in grade III two, in grade II three, and in grade I four, on the principle that mistakes in grade IV, the most difficult grade, were more to be expected than on grade I, etc. A low score, then, means good accuracy.

Records For Mistakes by Point Scaling

Points	*Men*	*Women*
0	2	5
1	6	13
2	8	10
3	10	10
4	6	7
5	10	8
6	11	2
7	3	1
8	2	4
9	2	2
10	3	6
11	1	3
12	1	2
13	2	0
14	2	0

The average score for the women by this method of point scaling was 4.64 and that of the men 6.48. In addition to the scores on the table, there were among the women a score of fifteen and one of sixteen, and among the men one each of 16, 17, 19, 24, 27, 29. Excluding these exceptional cases, we still have a lower average among the women, though the difference is not so marked. The average score for the women would then be 4.34, and for the men 5.13. These results imply that there were a larger number of men than women who made exceptionally bad records for accuracy, but that even with the exception of such cases, the average accuracy among the women was higher than among the men. This will be considered later under grouping of individuals.

Comparing the two groups for perfect or very good accuracy records, we find two men and four women with perfect accuracy records, five men and twelve women whose only mistake was a single one in the most difficult grade, six men and ten women whose only mistakes were in the most difficult grade, one man and no women whose only mistake was one in grade III and six men and five women whose only mistakes were in grades III and IV. There were then more very high accuracy records among the women.

A comparison of the two groups for ratios of times spent on more difficult grades as compared with more easy ones shows that both groups spent more time on the average on a more difficult grade as compared with an easier one. All of the ratios, except the ratio 2/1 (that is, of grade II as compared with grade I) were larger with the women than with the men. The differences were not marked, however, except in the ratios involving grade IV for which the women showed much larger ratios. The six ratios for the two groups and comparison of the two groups follow. The women, then, while taking more time than the men apportioned their time among the grades much as did the men with the exception of grade IV in which they took much more time as compared with the easier grades than did the men. Of the nine hundred individual ratio records (four hundred and fifty for the seventy-five men and four hundred and fifty for the seventy-five women) only sixty-seven or about 7% showed no gradation (no rise in time for a more difficult grade) or inverse gradation (less time on a more difficult grade). These sixty-seven instances are divided about equally among the two groups (thirty-five for the women and thirty-two for the men) and are distributed among the same numbers—one-third—of each group (twenty-five women and twenty-six men). Relatively few cases of inverse or lacking gradation occured between grades which were not successive. The largest number of inverse or lacking gradation with the women (fourteen) occurred between the two easiest grades, ratio 2/1, and the largest number with the men on ratios 4/3 (thirteen) and 3/2 (eleven).

Ratios	2/1	3/1	4/1	4/2	4/3	3/2
Women.................	1.55	2.43	4.06	2.68	1.70	1.56
Men....................	1.60	2.39	3.54	2.22	1.49	1.51

Ratios	2/1	3/1	4/1	4/2	4/3	3/2
Relation Women to Men.	1:1.03	1.01:1	1.14:1	1.20:1	1.14:1	1.03:1

Comparison of the two groups for variability of decision times within the grades shows for both groups lower variability on more difficult decisions with the exception of grade I. Since timing was only in seconds and the average time on grade I for men and women was respectively 2.76″ and 3.59″, the proportional variation in seconds on grade I is obviously not reliable. The women showed less variability in their decision times on all grades except the most difficult on which they were more variable than the men, as follows:

Proportional Variability	*Grade I*	*Grade II*	*Grade III*	*Grade IV*
Men.................	.370	.480	.395	.280
Women..............	.335	.436	.364	.307

To the end of finding correlations in the two groups of the numerous factors mentioned above, an attempt was made to separate the larger groups into smaller groups or decision types. Certain groups stand out at once as homogeneous.

Most conspicuous among the men is a very quick, ungraded (i.e. with low ratios or inverse ratios) inaccurate group. This group is also found among the women, but is not so numerous. It comprises 20% of the men and 12% of the women. While both of these groups are inaccurate, that of the men was more inaccurate than that of the women. A large percentage of the mistakes of the total group of men is due to this very quick group, all of the mistakes (seven) on grade I coming within this group. The average mistakes of this group as compared with total group of men were as follows:

	Grade I	Grade II	Grade III	Grade IV
Quick group........	.46	1.53	2.26	2.46
Total group........	.09	.42	1.32	2.18

They made 71% of the mistakes made by men on Grade II, whereas their number entitles them only to 20%, 34% of the mistakes on grade III and 22% of the mistakes on grade IV. The fact that this group approximates the average more closely on grade IV would seem to imply that even for fairly careful deliberation grade IV is difficult and that the difference between hasty and prolonged deliberation becomes less marked in results on this grade.

The quick women made the only mistake on grade I made by the women group, had 62% of the mistakes of the women group on grade II, 25% of the mistakes on grade III and 18% of the mistakes on grade IV, whereas their number entitles them to only 12% on each. The inaccuracy of the quick women, then, represents more of a divergence from the accuracy standard of the women group than does that of the men, though the absolute inaccuracy of the quick men was greater than that of women. These quick, inaccurate women averaged less quick in their decision times than the quick inaccurate men.

A very few of the men, (two, and possibly a third) making quick relatively ungraded decisions were able to make high accuracy records. This condition was not found among the women.

A second type of reaction was indicated in both sex groups, a slow accurate type. This was much more numerous among the women, comprising 24% of the women and only 14% of the men. The women of the group likewise averaged slower than the men and showed more extreme examples of slow decisions than the men. These slow groups are less homogeneous than the quick groups, but agree in being more accurate than the rest of the sex groups, though the women are more accurate than the men. They also average higher ratios. They divide into three general groups, the first characterized by high gradation of the ratios involving grade IV (that is, composed of those who took an excessive amount of the time on grade IV alone), the second characterized by high proportional variations (composed of people whose excessive time was due to occasional very large decision times on individual cards and not on whole grades), and the third composed of those who took a constantly excessive amount of time on all grades and on all cards within the grades. The first group was found only among the women (seven of the eighteen cases) and was an exceedingly accurate group (only one mistake on grade IV was made by the whole group). The condition of occasional very large decision times is not confined to this slow group and will be considered separately. One of the slow men and two of the slow women owed their excessive time to very slow decisions on the easiest grade alone. This excessively cautious group was accurate like the rest of the slow cases.

Since the quick inaccurate reaction is more numerous among the men and the slow accurate among the women, the remainder of the two sex groups (comprising approximately two thirds) were compared to see if the differences in the averages of the two sex groups are due to the existence of the two groups discussed above, quick inaccurate and slow accurate. This is found not to be the case. The remaining women are still slower and more accurate (except for grade III on which the two groups have practically the same accuracy) than the remaining men, though the groups approximate one another much more closely than do the general sex groups, as follows:

	Total Time	Grade I	Grade II	Grade III	Grade IV	Mistakes Per Individual	Grade I	Grade II	Grade III	Grade IV
2x3 Men......	124"	2.52	3.85	5.98	8.94		0	.18	1.14	2.12
2x3 Women...	151"	3.06	4.13	6.70	11.80		0	.10	1.16	1.61

Within this two-thirds group greater time still tends to be correlated with greater accuracy with both sexes, but for any given time women tend to be more accurate than men taking the same time.

The cases showing high proportional variability were principally of two kinds: the first consisted of cases who were fairly quick and irregular in their times, and the second those whose variability was due to occasional very delayed or blocked decisions. The quick irregular group (that is, those whose times averaged a given time which none of the times clustered closely around) was larger among the men and the blocked group much larger among the women. The fairly quick irregular group tended to be less accurate than both the whole group and the two-thirds group left after eliminating the quick inaccurate and slow accurate. Blocking seemed to be a fairly constant tendency among the women and blocked times shade into non-blocked times, while with the men the blocking tendency seems less characteristic and blocked decisions are more sharply differentiated in time from non-blocked decisions. Among both sex groups, cases characterized by blocking averaged more accurate than the whole group or the two-thirds group, though the most delayed separate decisions occurring in both sex groups were more inaccurate than the other decisions. The blocked decisions of the men showed a higher percentage of inaccuracies than those of the women. More extreme cases of blocking occurred among the women.

The constant decision time cases (those with low proportional variability) divide into two definite groups, a small accurate group, and an inaccurate group characterized by low ratios. The latter is considered an undeliberative group in which there was also no particular tendency towards impulsive action.

Considering decision times on separate cards apart from the individual cases to see if quicker times than the average of the individual or slower times were more likely to lead to inaccurate results, it was found that with both sexes more *wrong decisions* were correlated with decisions taking more than the average time of the individual for the grade than with decisions taking less, with the exception of grade IV on which slightly more of the mistakes were correlated with decisions taking less than the average. Both sexes had more mistakes correlated with average time decisions on grade IV than on any other grade, implying again the difficulty of grade IV. The most delayed decisions as compared with the average time of the individual had more than their quota of mistakes, as mentioned above, and the quickest fewer than their quota.

The women took a larger amount of time on the first few cards as compared with their average time than did the men. Though they thus adapted themselves less quickly to the new situation which the test implies, they adapted themselves more accurately, since they show fewer mistakes for the first few cards.

Considering the most accurate cases among the men and women, they are found to have slow time, high ratios and high porportional variations. The majority of the very inaccurate cases with both men and women belong to the very quick ungraded group, and the remainder show similar, if not such extreme, conditions—quick times and low ratios.

Of the cases showing lower gradation than the average on all ratios, a large majority with both sexes belong either to the very quick ungraded inaccurate group or to the slow, low variability, inaccurate group. Of the few cases remaining, two-thirds showed average accuracy and the other third poor accuracy.

Cases showing higher gradation than the average on all ratios were uniformly slow and accurate with both sexes.

The results then show fairly well marked decision groups, varying in their representation in the two sex groups, and some well marked decision tendencies and correlations of decision tendencies. The most marked groups are the very quick ungraded inaccurate group, most numerous among the men, and the slow graded accurate group, most numerous among the women, and the ungraded inaccurate group. Among the decision tendencies, most marked is the tendency toward frequent very delayed decision times among the women. This is an exaggeration of the general tendency toward slow decision times characteristic of most of the women group, contrasted with the tendency to very quick decisions among the men. The greater accuracy record of the women group throughout is striking. Among the correlations irrespective of the groups though compatible with the conditions found in them, are higher accuracy with slower times and higher ratios, and lower accuracy with quicker times and lower ratios.

BOOK REVIEW

HENRY C. LINK. *Employment Psychology*. The Macmillan Co., New York, 1920, p. xii +440.

A novel which attempts to tell a tale and preach a moral at the same time rarely succeeds in doing either particularly well. Any book taking major premises and attempting to draw minor premises, which shall result in true conclusions, labors under a dual handicap; its major premises are apt to be misstated, or its minor premises are liable to be only logical and not to work in practice. Just so, he who attempts to apply a science to the practice of human affairs has in advance our sympathy or our criticism (depending upon our temperaments and dispositions). If either the science or its applications are relatively new, these suspicions and sympathies are aggravated.

The applications of theoretical psychology to industrial practices are recent and its literature is correspondingly meager of quantity and on the whole bare of quality. Employment Psychology is Mr. Link's second book in this ambitious field. The book is divided into four parts: Part 1, Psychological Tests; Part 2, Trade Tests and other applications of employment psychology; Part 3, Selection and Retention, and Part 4, Conclusions.

Part 1—Psychological Tests—attempts to show the results of actual tests given to different types of workers, both in the office and in the shop, indicating the modifications that have been made to these tests to adapt them to differing circumstances. Difficulties are frankly considered and solutions of these difficulties indicated, so far as they have been discovered. The psychological background is fairly and on the whole clearly stated, and the results are tabulated with scientific impartiality. For the most part, the treatment is that of the psychological monograph but the usual heaviness of this procedure is enlivened by concrete circumstances and, for instance, by the treatment of general intelligence in a Platonic dialogue (Chapter 12), whose conclusions are that "There is no such thing as general intelligence, and that, if there were, it would be of little use to us in employment work because we are interested in specific abilities or *kinds* of intelligence, and not in degrees of intelligence, *per se*." The writer of the book believes that the application of psychological tests, in certain cases, makes it possible to discover facts otherwise discoverable only by the costly method of trial and error, but he does not offer them as a general panacea. In his own phrasing: "The scope of psychological tests has been discussed largely from the standpoint of their limitations." To be successful he believes that a test must be—a) carefully selected, b) based upon preliminary experiments upon a large group of workers similarly engaged, c) they must be checked by other and objective standards.

Under Part 2, the author's discussion of Job Analysis and the Vestibule School is clear, practical and to the point, and is in line with the best constructive thinking of the day. Part 3 is more general and theoretical although its brief chapters offer thoughtful reading to the employment manager.

The same comment also holds for the latter chapters of the book under the heading "Conclusions" in Part 4. Probably more theoretical but nevertheless important general principles of employment psychology are here discussed, such as the applicant's point of view, Employment Psychology, and the like. There is a considerable appendix and an index, the former outlining the tests upon which experiments described in the book are based.

Taken as a whole, this is a good book, and one particularly stimulating to the student of its problems. When one realizes the unstandardized and uncharted field that it attempts to cover, one inclines to superlative commendation. Certainly it is neither a text-book nor a book for leisure hours but what it purports to be, it is. Its introducer, Dr. Thorndike, says of it "an honest, impartial account of the use of psychological tests under working conditions in a representative industry." This it is.

In the present temper of the industrial mind, such a book will probably not be thoughtfully read by most of those who should read it—notably the employers of labor. By and large the employer has not yet come to place so high a value upon those who assemble his personnel as he does upon those who assemble his machines. Most employment managers are still clerks. It is also unfortunate, but true, that in times of labor scarcity, employers give anxious thought to these matters, while under conditions of labor surplus, selection is deemed to be easy and elaborate psychological tests only unnecessary frill and refinement.

ELIOTT FROST.

The following books and pamphlets have been received:*

Paul Bousfield. *The Elements of Practical Psycho-Analysis.* E. P. Dutton & Co., 1920. pp. 276.

A. T. Schofield, M. D. *The Mind of a Woman.* E. P. Dutton & Co. pp. 116.

W. Whately Smith. *The Foundations of Spiritualism.* E. P. Dutton & Co. New York. pp. 123.

Wm. Macpherson, D. C. *The Psychology of Persuasion.* E. P. Dutton and Co. New York. pp. 256.

Evelyn Dewey, Emily Child, Beardsley Ruml. *Methods and Results of Testing School Children.* E. P. Dutton & Co. New York. pp. 176.

Henri Pièron. *L'Annèe Psychologique.* Masson et Cie, Paris. 1920. pp. 547.

J. H. Berkowitz. *The Eyesight of School Children.* Washington, Government Printing Office. 1920. pp. 128.

Hermilio Valdizan and Honorio F. Delgado. *Revista de Psiquiatria y Disciplinas Conexas.* C. F. Southwell. Lima, Peru. 1920. pp. 103.

David E. Berg. *Personality Culture by College Faculties.* Institute for Public Service. New York City. 1920. pp. 127.

J. Alexander Gunn, M. A. *Bergson and his Philosophy.* E. P. Dutton & Co. New York. pp. 190.

Seventh Annual Conference on Educational Measurements. The Extension Division of Indiana University. 1920. pp. 85.

S. D. Porteus. *Condensed Guide to the Binet Tests.* Research Dept. of Training School at Vineland, N. J. 1920. pp. 39.

---

* Mention here does not preclude further comment.

JAMES SULLY. *My Life and Friends.* E. P. Dutton & Co. New York. pp. 344.

H. L. HOLLINGWORTH. *The Psychology of Functional Neuroses.* D. Appleton and Co. New York. 1920. pp. 259.

KNIGHT DUNLAP. *Mysticism, Freudianism and Scientific Psychology.* C. V. Mosby Co. St. Louis, Mo. 1920. pp. 173.

J. CROSBY CHAPMAN, B. A., Ph. D. *Trade Tests.* Henry Holt & Co. New York. 1921. pp. 435.

DR. J. SADGER, Vienna. *Sleep Walking and Moon Walking.* Nervous and Mental Disease Publishing Co. 1920. pp. 138.

J. B. SEARS. *The Boise Survey.* World Book Co. New York. 1920. pp. 290.

EDWARD J. KEMPF, M. D. *Psychopathology.* C. V. Mosby Co. St. Louis. 1920. pp. 762.

INDEX OF SUBJECTS

(The subject headings of all contributions except Book Reviews and Notes are printed in SMALL CAPITALS.)

Abnormal Psychology (Bridges) 108
ACADEMIC GRADES, CORRELATION WITH ARMY ALPHA TEST 289
ACADEMIC STANDING AND INTELLIGENCE STANDING 361
Adolescence, Education During (Mackie) 279
ALTERNATIVE INTELLIGENCE EXAMINATIONS, EQUALITY IN DIFFICULTY OF 283
ANALOGIES, STANDARDIZATION AND WEIGHTING OF 263
ANTICLIMAX AND CLIMAX OF PRESENTATION ON MEMORY, EFFECT OF 330
APPLIED PSYCHOLOGY OF NAMES 348
ARMY ALPHA INTELLIGENCE TEST CORRELATION WITH ACADEMIC GRADES 289
ARMY PROMOTION, PROPHESYING OF 73
ARMY TESTS, SOME RESULTS AND INFERENCES DERIVED FROM USE OF 59

Brightness and Dullness in Children (Woodrow) 279

CLIMAX AND ANTICLIMAX, EFFECT OF ORDER OF PRESENTATION ON MEMORY 330
CLINICAL PSYCHOLOGY AND THE DEGREE OF PH. D. 88
CLINICAL STUDY AS METHOD IN EXPERIMENTAL EDUCATION 126
CO-EFFICIENTS OF CORRELATION BY RANK DIFFERENCE METHOD, TABLES OF 115
COLLEGE GRADUATES, PSYCHOLOGICAL EXAMINING IN SOCIAL SERVICE AND EDUCATION BY 207
COLLEGE WOMEN, PSYCHOLOGICAL *TESTS AS DIAGNOSTIC* VOCATIONAL APTITUDES IN 30
COLORED AND WHITE COLLEGE STUDENTS, STUDY OF, BY STANFORD BINET SCALE 316
COMMON SYNAESTHESIA OF MUSIC 342
COMPLETION, PICTURE 310
COMPLETION TEST, COMPARISON OF TWO METHODS OF GIVING NUMBER SERIES 346
CONSTANT ERROR IN PSYCHOLOGICAL RATINGS 25
CORRELATION, ARMY ALPHA AND ACADEMIC GRADES, 289; BY RANK DIFFERENCE METHOD, TABLES OF 115

DECISION STUDY OF YOUNG MEN AND WOMEN 364
Dullness and, Brightness in Children (Woodrow) 279

Education during Adolescence (Mackie) 279
EDUCATION, CLINICAL STUDY AS A METHOD IN EXPERIMENTAL, 126; PSYCHOLOGICAL EXAMINING BY COLLEGE GRADUATES IN 207
EFFECTS OF VARIOUS MODES OF PUBLIC READING 162
EMOTIONS, *FIRST REVISION* GROUP SCALE FOR INVESTIGATING 97
Employment Psychology (Link) 375
EMPLOYMENT PSYCHOLOGY, FURTHER DEVELOPMENT OF, 306; IN PUBLIC INDUSTRY 1
ENVIRONMENT, EFFECT OF POOR, ON TESTS OF INTELLIGENCE 91

EXAMINATIONS, EQUALITY IN DIFFICULTY OF ALTERNATIVE INTELLIGENCE 283
EXPERIMENTAL EDUCATION, CLINICAL STUDY AS A METHOD OF 126

Factory Labor, Turnover of (Slichter) 278
FATIGUE, RACIAL DIFFERENCES IN 235
FEEBLEMINDED NEGRO, PHENOMENAL MEMORIZING BY 202

GENERAL INTELLIGENCE, DETERMINED BY WEAKEST ESSENTIAL ELEMENT 155
GROUP SCALE, FOR INVESTIGATING EMOTIONS, FIRST REVISION, TENTATIVE NORMS 97
Guide for Stanford Revision Binet-Simon Intelligence Tests (Terman) 281

Health, Towards Racial (March) 276

INADEQUATE SCHOOLING, EFFECT OF, ON RESULTS OF TESTS OF INTELLIGENCE 91
Industrial Education (Dooley) 110
INDUSTRY, A NEW APPLICATION OF PSYCHOLOGY TO, 245; WHAT IT WANTS AND DOES NOT WANT FROM PSYCHOLOGISTS 18
INFLUENCE OF (a) INADEQUATE SCHOOLING (b) POOR ENVIRONMENT ON RESULTS OF TESTS OF INTELLIGENCE 91
INTELLIGENCE, GENERAL, AS DETERMINED BY WEAKEST ESSENTIAL ELEMENT, 155; INTELLIGENCE AND SENESCENCE, 219; EXAMINATIONS, EQUALITY IN DIFFICULTY OF, 283; TESTS AND ACADEMIC STANDING, 361; TESTS, INADEQUATE SCHOOLING AND POOR ENVIRONMENT, 91; OF WHITE AND COLORED STUDENTS BY STANFORD REVISION OF BINET 316

LABOR, TURNOVER OF FACTORY 278
LEARNING TEST, TENTATIVE NORMS IN RATIONAL 250

Management, Psychology of (Gilbreth) 274
MEASUREMENT, SHORT POINT SCALE FOR MENTAL 258
Measuring Results of Teaching (Monroe) 281
MEMORIZING, PHENOMENAL, BY FEEBLEMINDED NEGRO 202
MEMORY, EFFECT OF CLIMAX AND ANTICLIMAX ORDER OF PRESENTATION ON 330
MEMORY TESTS, SIGNIFICANT RESPONSES IN CERTAIN 142
MENTAL FATIGUE, RACIAL DIFFERENCES IN, 235; MEASUREMENT, SHORT POINT SCALE FOR, 258; TESTS, APPLICABILITY OF, TO PERSONS OVER FIFTY YEARS OF AGE, 39; TESTS AND SELECTION OF MILL WORKERS, 339
METHOD IN EXPERIMENTAL EDUCATION, CLINICAL STUDY AS A 126
MILL WORKERS, SELECTION BY MENTAL TESTS 339
MINOR STUDIES, PSYCHOLOGICAL LABORATORY, INDIANA UNIVERSITY 91
MUSIC, COMMON SYNAESTHESIA OF 342
Musical Talent, Psychology of (Seashore) 105
Nationality and Internationalism, Psychology of (Pillsbury) 274
NEGRO, CASE OF PHENOMENAL MEMORIZING BY FEEBLEMINDED 202
NAMES, APPLIED PSYCHOLOGY OF 348
NORMS, TENTATIVE, GROUP SCALE FOR INVESTIGATING EMOTIONS, 97; TENTATIVE IN RATIONAL LEARNING TEST 250
Notes and News:
  National Academy of Science, 113; National Research Council, 113; State Psychiatric Clinic of New Jersey, 113; Work of Psychologists on Job-Analysis, Man-Analysis and coordination of these 113, 114

Number Series Completion Test, Comparison of Two Methods of Giving 346

Picture Completion Test 310
Point Scale, Short, for Mental Measurement 258
Prophesying Army Promotion 73
Psychical Science (Crawford) 280
Psychoanalysis, History, Theory, Practice (Tridon) 277
Psychological Examining in Social Service and Education, 207; Ratings, Constant Error in, 25; Tests as Diagnostic of Vocational Aptitudes in College Women 30
Psychologists, What Industry Wants and Does Not Want from 18
Psychology, Applied, of Names, 348; Bake Bread, Should, 294; Further Development of Employment, 306; of Management (Gilbreth), 274; of Musical Talent (Seashore), 105; of Nationality and Internationalism (Pillsbury), 274; New Application of, to Industry 245

Racial Differences in Mental Fatigue, 235; Health (March) 276
Rank Difference Method, Tables 115
Ratings, Constant Error in Psychological 25
Results and Inferences, Derived from Use of Army Tests at University of Minnesota 59

Scale for Investigating of Emotions, First Revision Tentative Norms, 97; for Mental Measurement, Short Point 258
School Statistics and Publicity (Alexander) 109
Selection of Mill Workers by Mental Tests, 339; Vocational, for Specialized Tasks 186
Social Service Work, Psychological Examining by College Graduates in 207
Standardization, and Weighting of 200 Analogies 263
Stanford Revision of Binet-Simon Scale with White and Colored College Students, Results of 316
Senescence and Intelligence 219
Synaesthesia, Common, of Music 342

Tables to Facilitate Computation of Coefficients of Correlation by Rank Difference Method 115
Talent, Psychology of Musical (Seashore) 105
Tasks, Vocational Selection for Specialized 186
Teaching, Measuring Results of (Monroe) 281
Tests, Academic Standing and Intelligence, 361; Diagnostic, of Vocational Aptitudes in College Women, 30; Intelligence, Inadequate Schooling and Poor Environment, 91; Norms in Rational Learning, 250; Number Series Completion, Comparison of Two Methods of Giving, 346; Selection of Mill Workers by Mental, 339; Significant Responses in Certain Memory 142
Turnover of Factory Labor (Slichter) 278

Vocational Aptitudes, in College Women 30
Vocational Selection for Specialized Tasks 186

Weighting and Standardization of 200 Analogies 263
White and Colored College Students, Comparative Intelligence of, by Stanford Binet Scale 316
Workers, Mill, Selection by Mental Tests 339

INDEX OF AUTHORS

(The names of authors of original contributions are printed in SMALL CAPITALS.)

ADAMS, HENRY FOSTER 330
Alexander, Carter 109
ANDERSON, JOHN E. 346
ARPS, G. F. 289

BAKER, H. J. 361
BARRETT, MYRON F. 278
BATES, MARJORY 108
BAXTER, MILDRED FRANCES 207
BEESON, M. F. 219
BLANCHARD, PHYLLIS 276, 277
BRIDGES, JAMES W. 108, 258
BURTT, H. E. 1, 289
BYRD, HIRAM 202

CHAMBERS, O. R. 97
COE, GEORGE ALLEN 109
Crawford, W. J. 280

DERRICK, S. M. 316
DOLL, EDGAR A. 88
Dooley, William H. 110

FOSTER, JOSEPHINE CURTIS 39, 142
FREEMAN, FRANK N. 126
FROST, ELIOTT 18, 294, 375

GARTH, THOMAS R. 235
GAULT, ROBERT H 310
GEISSLER, L. R. 279, 280, 281
GIBSON, SYBIL MARSH 364
Gilbreth, L. M. 274
GORDON, EDNA 361
GROETZINGER, MARIE 348

IRLE, K. W. 73

KOHS, S. C. 73

LI, CHI 274
LINK, HENRY C. 245, 306

Mackie, Ransom A. 279
March, Norah H. 276
MARCUS, LAWRENCE 186
MILLIKEN, HARRY E. 110
Monroe, Walter Scott 281
MOORE, BRUCE V. 155
MUDGE, E. LEIGH 342
MURRAY, ELSIE 30

OGDEN, R. M. 105
OTIS, ARTHUR S. 339

PETERSON, JOSEPH 250
Pillsbury, W. B. 274
PINTNER, RUDOLPH 263
PRESSEY, LUELLA WINIFRED 91
PRESSEY, S. L. 97

REEDY, ESTHER 258
RENSHAW, SAMUEL 263
ROBACK, A. A. 348

Seashore, C. E. 105
Slichter, Sumner H. 278
SPROWLS, J. W. 274

TAYLOR, GRACE A. 39
Terman, Lewis M. 281
THORNDIKE, EDWARD L. 25, 283
Tridon, André 277

VAN WAGENEN, M. J. 59
Woodrow, Herbert 279
WOOLBERT, CHARLES H. 162